

# VIRTUAL STRINGS

VLADIMIR TURAEV

**ABSTRACT.** A virtual string is a scheme of self-intersections of a closed curve on a surface. We introduce virtual strings and study their geometric properties and homotopy invariants. We also discuss connections between virtual strings, Gauss words, and virtual knots.

## Contents

1. Introduction
2. Generalities on virtual strings
3. Polynomial  $u$
4. Geometric realization of virtual strings
5. Combinatorics of closed curves on the 2-sphere
6. Based skew-symmetric matrices
7. Based matrices of strings
8. Lie cobracket for strings
9. Virtual strings versus virtual knots
10. Proof of Theorem 9.3.1
11. Algebras and groups associated with strings
12. Open questions

## 1. INTRODUCTION

A virtual string is a scheme of self-intersections of a generic oriented closed curve on an oriented surface. More precisely, a virtual string of rank  $m \geq 0$  is an oriented circle with  $2m$  distinguished points partitioned into  $m$  ordered pairs. These  $m$  ordered pairs of points are called arrows of the virtual string. An example of a virtual string of rank 3 is shown on Figure 1 where the arrows are represented by geometric vectors.

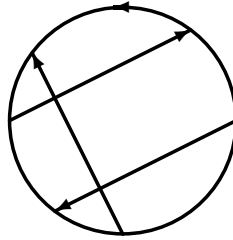


FIGURE 1. A virtual string of rank 3

A (generic oriented) closed curve on an oriented surface gives rise to an “underlying” virtual string whose arrows correspond to the self-crossings of the curve. The usual homotopy of curves suggests a notion of homotopy for strings. The main objective of the theory of virtual strings is a study (and eventually classification) of their homotopy classes. To this end, we introduce certain homotopy invariants of virtual strings, specifically a one-variable polynomial  $u$  and a so-called based matrix.

In analogy with the theory of knots we define a class of slice virtual strings. A string is slice if it underlies a closed curve on a closed surface  $\Sigma$  which is contractible in a handlebody bounded by  $\Sigma$ . We formulate obstructions to the sliceness of a string in terms of the polynomial  $u$  and the based matrix.

We introduce a natural Lie cobracket in the free abelian group generated by the homotopy classes of strings. Dually, the abelian group of  $\mathbb{Z}$ -valued homotopy invariants of strings becomes a Lie algebra. This

Lie algebra gives rise to further algebraic objects including a Hopf algebra structure in the (commutative) polynomial algebra generated by the homotopy classes of strings.

The notion of a virtual string should be compared with the one of a Gauss word. A Gauss word (or a Gauss code) is a sequence of letters in a finite alphabet in which all letters of the alphabet occur exactly twice. The sequence is considered up to circular permutations. Gauss observed that a closed curve on  $\mathbb{R}^2$  gives rise to a Gauss word. Namely, label the self-crossings of the curve by different letters and write down these letters in the order of their appearance when one traverses the whole curve beginning at a generic point. The resulting word is well defined up to circular permutations (and the choice of letters). A similar procedure applies to virtual strings where instead of crossings one should label the arrows. Thus, each virtual string gives rise to a Gauss word. In fact, the notion of a Gauss word is equivalent to the notion of a virtual string with the order in the distinguished pairs of points forgotten.

Gauss introduced his sequences of letters in an attempt to give a combinatorial formulation of closed curves on  $\mathbb{R}^2$ . We discuss another relevant ingredient: a bipartition of the alphabet. As an application of the theory of virtual strings, we obtain a complete combinatorial description of closed curves on the 2-sphere in terms of Gauss words and bipartitions. This description extends the theorem of P. Rosenstiehl [Ro] characterizing the Gauss words realizable by closed curves on  $\mathbb{R}^2$ .

Virtual strings are closely related to virtual knots introduced by L. Kauffman [Ka]. They can be defined in terms of so-called arrow diagrams which are just virtual strings whose arrows are provided with signs  $+$  or  $-$ . Virtual knots are arrow diagrams considered up to several moves, induced by the Reidemeister moves on ordinary knot diagrams. The term “virtual knots” suggested to us the term virtual strings.

Forgetting the signs of the arrows transforms an arrow diagram into a virtual string. This gives a map from the set of virtual knots into the set of homotopy classes of virtual strings. We give a more elaborate construction which associates with each virtual knot a polynomial expression in virtual strings with coefficients in the ring  $\mathbb{Q}[z]$ . This leads to an isomorphism between a skein algebra of virtual knots, defined below, and a polynomial algebra generated by the homotopy classes of strings.

The organization of the paper should be clear from the Contents above.

## 2. GENERALITIES ON VIRTUAL STRINGS

**2.1. Definitions.** We give here a formal definition of a virtual string. For an integer  $m \geq 0$ , a *virtual string*  $\alpha$  of rank  $m$  (or briefly a *string*) is an oriented circle,  $S$ , called the *core circle* of  $\alpha$ , and a distinguished set of  $2m$  distinct points of  $S$  partitioned into  $m$  ordered pairs. We call these  $m$  ordered pairs of points the *arrows* of  $\alpha$ . The set of arrows of  $\alpha$  is denoted  $\text{arr}(\alpha)$ . The endpoints  $a, b \in S$  of an arrow  $(a, b) \in \text{arr}(\alpha)$  are called its *tail* and *head*, respectively. The  $2m$  distinguished points of  $S$  are called the *endpoints* of  $\alpha$ .

The string formed by an oriented circle and an empty set of arrows is called a *trivial virtual string*. An example of a virtual string of rank 3 is shown on Figure 1.

By a *homeomorphism* of two virtual strings, we mean an orientation-preserving homeomorphism of the core circles transforming the set of arrows of the first string onto the set of arrows of the second string. Two virtual strings are *homeomorphic* if they are related by a homeomorphism. Clearly, homeomorphic strings have the same rank.

By abuse of language, the homeomorphism classes of virtual strings will be also called virtual strings.

**2.2. From curves to strings.** By a surface, we mean a smooth oriented 2-dimensional manifold. By a *closed curve* on a surface  $\Sigma$ , we mean a generic smooth immersion  $\omega$  of an oriented circle  $S$  into  $\Sigma$ . Recall that a smooth map  $S \rightarrow \Sigma$  is an *immersion* if its differential is non-zero at all points of  $S$ . An immersion  $\omega : S \rightarrow \Sigma$  is *generic* if  $\#(\omega^{-1}(x)) \leq 2$  for all  $x \in \Sigma$ , the set  $\{x \in \Sigma \mid \#(\omega^{-1}(x)) = 2\}$  is finite, and all its points are transverse intersections of two branches. Here and below the symbol  $\#(A)$  denotes the cardinality of a set  $A$ . The points  $x \in \Sigma$  such that  $\#(\omega^{-1}(x)) = 2$  are called *double points* or *crossings* of  $\omega$ .

A closed curve  $\omega : S \rightarrow \Sigma$  gives rise to an *underlying virtual string*  $\alpha_\omega$ . The core circle of  $\alpha_\omega$  is  $S$  and the arrows of  $\alpha_\omega$  are all ordered pairs  $a, b \in S$  such that  $\omega(a) = \omega(b)$  and the pair (a positive tangent vector of  $\omega$  at  $a$ , a positive tangent vector of  $\omega$  at  $b$ ) is a positive basis in the tangent space of  $\omega(a)$ . For instance, the underlying string of a simple closed curve on  $\Sigma$  is a trivial virtual string.

We say that a virtual string is *realized* by a closed curve  $\omega : S \rightarrow \Sigma$  if it is homeomorphic to  $\alpha_\omega$ . As we shall see below, every virtual string can be realized by a closed curve on a surface.

**2.3. Homotopy of strings.** The usual homotopy of closed curves on a surface suggests to introduce a relation of homotopy for virtual strings. Observe first that two homotopic curves on a surface can be related by a finite sequence of the following “elementary” moves (and the inverse moves):

- (a) a local move adding a small curl to the curve;
- (b) a local move pushing a branch of the curve across another branch and creating two new double points;
- (c) a local move pushing a branch of the curve across a double point;
- (d) ambient isotopy in the surface.

The move (a) has two forms  $(a)^+$  and  $(a)^-$  depending on whether the curl lies on the left or the right of the curve where the left and the right are determined by the direction of the curve and the orientation of the surface. Considered up to ambient isotopy, the move (b) has three forms depending on the direction of the two branches. Similarly, considered up to ambient isotopy, the move (c) has two forms  $(c)^+$  and  $(c)^-$  depending on the direction of the branches. Using the standard braid generators  $\sigma_1, \sigma_2$  on 3 strands we can encode this move as  $\sigma_1\sigma_2\sigma_1 \mapsto \sigma_2\sigma_1\sigma_2$  where the over/undercrossing information is forgotten. The moves  $(c)^+$  and  $(c)^-$  are obtained by directing (before and after the move) the first and third strands up and the second strand up or down, respectively. It is easy to see that  $(c)^+, (c)^-$  can be obtained from each other using ambient isotopy, moves (b), and inverses to (b). Similarly, the moves  $(a)^+, (a)^-$  can be obtained from each other using ambient isotopy, moves (b),  $(c)^-$ , and inverses to (b). Thus the moves  $(a)^-, (b), (c)^-$  generate all the other moves.

It is clear that ambient isotopy of a closed curve does not change the underlying virtual string. We now describe the analogues for virtual strings of the moves  $(a)^-, (b), (c)^-$ . In this description and in the sequel, by an *arc* on an oriented circle  $S$  we mean an *embedded arc* on  $S$ . The orientation of  $S$  induces an orientation of all arcs on  $S$ . For two distinct points  $a, b \in S$ , we write  $ab$  for the unique oriented arc in  $S$  which begins in  $a$  and terminates in  $b$ . Clearly,  $S = ab \cup ba$  and  $ab \cap ba = \{a, b\}$ .

Let  $\alpha$  be a virtual string with core circle  $S$ . Pick two distinct points  $a, b \in S$  such that the arc  $ab \subset S$  is disjoint from the set of endpoints of  $\alpha$ . The move  $(a)_s$ , where  $s$  stands for “string”, adds to  $\alpha$  the pair  $(a, b)$ . This amounts to attaching a small arrow to  $S$  such that the arc in  $S$  leading from its tail to its head is disjoint from the endpoints of  $\alpha$ . The move  $(b)_s$  acts on  $\alpha$  as follows. Pick two arcs on  $S$  disjoint from each other and from the endpoints of  $\alpha$ . Let  $a, a'$  be the endpoints of the first arc (in an arbitrary order) and  $b, b'$  be the endpoints of the second arc. The move  $(b)_s$  adds to  $\alpha$  two arrows  $(a, b)$  and  $(b', a')$ . (This move has four forms depending on the two possible choices for  $a$  and two possible choices for  $b$ . However, two of these forms of  $(b)_s$  are equivalent.) The move  $(c)_s$  applies to  $\alpha$  when  $\alpha$  has three arrows  $(a^+, b), (b^+, c), (c^+, a)$  where  $a, a^+, b, b^+, c, c^+ \in S$  such that the arcs  $aa^+, bb^+, cc^+$  are disjoint from each other and from the other endpoints of  $\alpha$ . The move  $(c)_s$  replaces the arrows  $(a^+, b), (b^+, c), (c^+, a)$  with the arrows  $(a, b^+), (b, c^+), (c, a^+)$ .

We say that two virtual strings are *homotopic* if they can be related by a finite sequence of homeomorphisms, the *homotopy moves*  $(a)_s, (b)_s, (c)_s$ , and the inverse moves. A virtual string homotopic to a trivial virtual string is said to be *homotopically trivial*.

It is clear from what was said above that the underlying virtual strings of homotopic closed curves on a surface are themselves homotopic.

**2.4. Transformations of strings.** For a string  $\alpha$ , we define the *opposite string*  $\alpha^-$  to be  $\alpha$  with opposite orientation on the core circle. The *inverse string*  $\bar{\alpha}$  is obtained from  $\alpha$  by reversing all its arrows. On the level of closed curves on surfaces, these two transformations correspond to traversing the same curve in the opposite direction and to inverting the orientation of the ambient surface, respectively. If two strings are homotopic, then their opposite (resp. inverse) strings are homotopic.

One can raise a number of questions concerning the transformations  $\alpha \mapsto \alpha^-, \alpha \mapsto \bar{\alpha}, \alpha \mapsto \bar{\alpha}^-$ . For instance, one can ask whether there is a string  $\alpha$  that is not homotopic to  $\alpha^-$  (resp. to  $\bar{\alpha}, \bar{\alpha}^-$ ). Below we will answer this question in the positive.

A virtual string  $\alpha$  with core circle  $S$  is a *product* of virtual strings  $\alpha_1$  and  $\alpha_2$  if there are disjoint arcs  $a_1b_1, a_2b_2 \subset S$  such that each arrow of  $\alpha$  has both endpoints on either  $a_1b_1$  or on  $a_2b_2$  and the string formed by  $S$  and the arrows of  $\alpha$  with endpoints on  $a_ib_i$  is homeomorphic to  $\alpha_i$  for  $i = 1, 2$ . One can ask whether the product is a well-defined operation on strings (at least up to homotopy) and whether it is commutative. Below we will answer these questions in the negative.

**2.5. Geometric invariants of strings.** We define four geometric characteristics of strings: the genus, the homotopy genus, the slice genus, and the homotopy rank. For a string  $\alpha$ , its *genus*  $g(\alpha)$  is the minimal integer  $g \geq 0$  such that  $\alpha$  can be realized by a closed curve on a surface of genus  $g$ . The *homotopy genus*  $hg(\alpha)$  is the minimal integer  $g \geq 0$  such that  $\alpha$  is homotopic to a string of genus  $g$ . The *homotopy rank*  $hr(\alpha)$  is the minimal integer  $m \geq 0$  such that  $\alpha$  is homotopic to a string of rank  $m$ . For example, if  $\alpha$  is a trivial string, then  $g(\alpha) = hg(\alpha) = hr(\alpha) = 0$ . It is clear that the homotopy genus and the homotopy rank are homotopy invariants of strings. Below we compute the genus explicitly and show that it is not a homotopy invariant.

For  $k \geq 0$  denote by  $\Sigma_k$  a compact oriented surface of genus  $k$  bounded by a circle. The *slice genus*  $sg(\alpha)$  of a string  $\alpha$  is the minimal integer  $k \geq 0$  such that there is a handlebody  $H$  (of a certain genus) and a map  $\Omega : \Sigma_k \rightarrow H$  such that  $\Omega(\partial\Sigma_k) \subset \partial H$  and the map  $\Omega|_{\partial\Sigma_k} : \partial\Sigma_k \rightarrow \partial H$  is a (generic) closed curve on  $\partial H$  representing  $\alpha$ . The existence of such  $k$  follows from the fact that any loop on a closed surface becomes homologically trivial in a certain handlebody bounded by this surface.

If  $sg(\alpha) = 0$ , then we say that  $\alpha$  is *slice*. Thus, a string is slice if it can be realized on a closed surface by a curve contractible in a handlebody bounded by this surface. For example, a trivial string is slice.

The genus, the homotopy genus, the slice genus, and the homotopy rank of a string are preserved under the transformations  $\alpha \mapsto \alpha^-$ ,  $\alpha \mapsto \bar{\alpha}$ .

**2.6. Encoding of strings.** There are two simple methods allowing to encode virtual strings in a compact way. Although we do not use these methods in this paper, we briefly describe them for completeness.

(1) Consider a finite set  $E$  consisting of  $m$  elements and its disjoint copy  $E^+ = \{x^+ | x \in E\}$ . Let  $y_1, y_2, \dots, y_{2m}$  be a sequence of elements of the set  $E \cup E^+$  in which every element appears exactly once. (Such a sequence determines a total order in  $E \cup E^+$  and vice versa.) The sequence  $y_1, y_2, \dots, y_{2m}$  defines a string of rank  $m$  whose underlying circle is  $S = \mathbb{R} \cup \{\infty\}$  with right-handed orientation on  $\mathbb{R}$  and whose  $m$  arrows are the pairs  $(a, b)$  such that  $a, b \in \{1, 2, \dots, 2m\} \subset S$ ,  $y_a \in E$ , and  $y_b = y_a^+ \in E^+$ . Any string can be encoded in this way. For instance, the string drawn in Figure 1 is encoded by the sequence  $x^+, y, z^+, x, z, y^+$  where  $E = \{x, y, z\}$ .

(2) By Section 2.2, virtual strings can be encoded by closed curves on surfaces. This has an extension similar to Kauffman's graphical encoding of virtual knots in [Ka]. Namely, consider a (generic) closed curve on a surface and suppose that some of its crossings are marked as "virtual". Take the string of this curve as in Section 2.2 and forget all its arrows corresponding to virtual crossings. It is easy to see that every virtual string can be obtained in this way from a closed curve in  $\mathbb{R}^2$  with virtual crossings. This yields a graphical encoding of strings by plane curves with virtual crossings. The relation of homotopy for strings has a simple description in this language: it is generated by the moves shown in [Ka], Figure 2 (where the over/undercrossing information should be forgotten).

**2.7. Remarks.** 1. We can point out certain classes of closed curves on surfaces whose underlying virtual strings are homotopically trivial. Since all closed curves on  $S^2$  are contractible, their underlying strings are homotopically trivial. Therefore the same is true for closed curves on any subsurface of  $S^2$ , i.e., on any surface of genus 0. In particular, all closed curves on an annulus have homotopically trivial underlying strings. Since each closed curve on a torus can be deformed into an annulus, its underlying string is homotopically trivial. The same holds for closed curves on a torus with holes.

2. The move  $(a)_s$  has a version  $(a)_s^+$  which is defined as  $(a)_s$  above but adds the arrow  $(b, a)$  rather than  $(a, b)$ . This move underlies the move  $(a)^+$  on closed curves. The move  $(a)_s^+$  preserves the homotopy class of a string. Indeed, it can be expressed as a composition of  $(b)_s$ ,  $(c)_s$ , and an inverse to  $(a)_s$ .

3. The move  $(c)_s$  has a version  $(c)_s^+$  which applies to a string when it has three arrows  $(a, b), (a^+, c), (b^+, c^+)$  such that the arcs  $aa^+, bb^+, cc^+$  are disjoint from each other and from the other endpoints of the string. The move  $(c)_s^+$  replaces these three arrows with the arrows  $(a^+, b^+), (a, c^+), (b, c)$ . This move underlies the move  $(c)^+$  on closed curves. The move  $(c)_s^+$  can be expressed as a composition of  $(c)_s^-$  and  $(b)_s$ .

4. For any virtual string  $\alpha$ , its appropriate product with  $\bar{\alpha}^-$  is slice. Indeed, let  $S$  be the core circle of  $\alpha$  and let  $ab \subset S$  be an arc containing all the endpoints of  $\alpha$ . Let  $(\alpha', S', a'b' \subset S')$  be a disjoint copy of the triple  $(\alpha, S, ab)$ . Consider the circle  $S'' = (ab \cup a'b')/a = a', b = b'$  and provide it with the orientation extending the one on  $ab$ . The arrows of  $\alpha$  and  $\bar{\alpha}'$  are attached to  $ab \cup a'b'$  and form in this way a virtual string,  $\alpha''$ , with core circle  $S''$ . It is clear that  $\alpha''$  is a product of  $\alpha$  with  $\bar{\alpha}^-$ . We claim that  $\alpha''$  is slice. To see this, represent  $\alpha$  by a closed curve  $\omega : S \rightarrow \Sigma$  on a surface  $\Sigma$ . The map  $\omega$  transforms  $S - ab$  onto an embedded arc in  $\Sigma$  disjoint from the rest of the curve. Let  $D \subset \Sigma$  be a 2-disc such that  $D \cap \omega(S) = \omega(S - ab)$

and  $\partial D \cap \omega(S) = \{\omega(a), \omega(b)\}$ . The 3-manifold  $H = (\Sigma - \text{Int } D) \times [0, 1]$  is a handlebody. The four paths  $\omega(ab) \times 0, \omega(ab) \times 1, \omega(a) \times [0, 1], \omega(b) \times [0, 1]$  form a closed curve on  $\partial H$  realizing  $\alpha''$  and contactible in  $H$ .

5. Replacing the circle in the definition of a string by an oriented one-dimensional manifold  $X$  we obtain a *virtual string with core manifold  $X$* . The definitions and results of this paper can be extended to such strings with appropriate changes. Of special interest are strings with core manifold  $X = [0, 1]$ ; we call them *open strings*. In this context it is natural to call virtual strings with core manifold a circle *closed strings*. Gluing the endpoints of  $X = [0, 1]$ , we can transform an open string into a closed string, its *closure*. Open strings underlie (generic) paths on a surface with endpoints on the boundary. Open strings can be multiplied in the obvious way.

**2.8. Exercise.** Verify that all virtual strings of rank  $\leq 2$  are homotopically trivial.

### 3. POLYNOMIAL $u$

**3.1. Invariants  $\{u_k\}_k$ .** Let  $\alpha$  be a virtual string with core circle  $S$ . Each arrow  $e = (a, b) \in \text{arr}(\alpha)$  splits  $S$  into two arcs  $ab$  and  $ba$ . We say that an arrow  $f = (c, d)$  of  $\alpha$  (distinct from  $e$ ) *links*  $e$  if one of its endpoints lies on  $ab$  and the other one lies on  $ba$ . More precisely,  $f = (c, d)$  links  $e$  *positively* (resp. *negatively*) if  $c \in ab, d \in ba$  (respectively, if  $c \in ba, d \in ab$ ). If  $f$  does not link  $e$ , then  $e$  and  $f$  are *unlinked*. Let  $n(e) \in \mathbb{Z}$  be the algebraic number of arrows of  $\alpha$  linking  $e$ , i.e., the number of arrows of  $\alpha$  linking  $e$  positively minus the number of arrows of  $\alpha$  linking  $e$  negatively.

It is easy to trace the behaviour of  $n(e)$  under the homotopy moves  $(a)_s, (b)_s, (c)_s$  on  $\alpha$ . The move  $(a)_s$  adds an arrow  $e_0$  with  $n(e_0) = 0$  and keeps  $n(e)$  for all other arrows. The move  $(b)_s$  adds two arrows  $e_1, e_2$  with  $n(e_1) = -n(e_2)$  and keeps  $n(e)$  for all other arrows. Consider the move  $(c)_s$  and use the notation of Section 2.3. It is obvious that for all arrows  $e$  preserved under the move, the number  $n(e)$  is also preserved. Each arrow  $e = (a^+, b), (b^+, c), (c^+, a)$  occurring before the move gives rise to an arrow  $e' = (a, b^+), (b, c^+), (c, a^+)$ , respectively, occurring after the move. We claim that  $n(e) = n(e')$ . Consider for concreteness  $e = (a^+, b)$ . Note that the points  $c, c^+$  lie either on  $ab$  or on  $ba$ . Suppose that  $c, c^+ \in ab$ . Then the arrows  $(b^+, c)$  and  $(c^+, a)$  contribute 1 and  $-1$  to  $n(e)$ , respectively, while the corresponding arrows  $(b, c^+)$  and  $(c, a^+)$  contribute 0 to  $n(e')$ . All other arrows contribute the same to  $n(e)$  and  $n(e')$ . Hence  $n(e) = n(e')$ . If  $c, c^+ \in ba$ , then the arrows  $(b^+, c)$  and  $(c^+, a)$  contribute 0 to  $n(e)$  while the corresponding arrows  $(b, c^+)$  and  $(c, a^+)$  contribute  $-1$  and  $1$  to  $n(e')$ , respectively. All other arrows contribute the same to  $n(e)$  and  $n(e')$ . Hence  $n(e) = n(e')$ .

For an integer  $k \geq 1$ , set

$$u_k(\alpha) = \#\{e \in \text{arr}(\alpha) \mid n(e) = k\} - \#\{e \in \text{arr}(\alpha) \mid n(e) = -k\} \in \mathbb{Z}.$$

It is clear from what was said above that  $u_k(\alpha)$  is preserved under the moves  $(a)_s, (b)_s, (c)_s$ . In other words,  $u_k(\alpha)$  is a homotopy invariant of  $\alpha$ . Clearly,  $u_k(\alpha) = 0$  for all  $k$  greater than or equal to the rank of  $\alpha$ . If  $\alpha$  is homotopically trivial, then  $u_k(\alpha) = 0$  for all  $k \geq 1$ .

**3.2. Polynomial  $u(\alpha)$ .** We can combine the invariants  $u_k$  of a virtual string  $\alpha$  into a polynomial

$$u(\alpha) = \sum_{k \geq 1} u_k(\alpha) t^k$$

where  $t$  is a variable. The free term of this polynomial is always 0 and its degree is bounded from above by  $m - 1$  where  $m$  is the rank of  $\alpha$ . This polynomial is a homotopy invariant of  $\alpha$ . If  $\alpha$  is homotopically trivial, then  $u(\alpha) = 0$ . (The converse is not true, as we shall see below.) The polynomial  $u(\alpha)$  yields an estimate for the homotopy rank  $hr(\alpha)$  of  $\alpha$  defined in Section 2.5:

$$(3.2.1) \quad hr(\alpha) \geq \deg u(\alpha) + 1.$$

We can rewrite  $u(\alpha)$  as follows:

$$(3.2.2) \quad u(\alpha) = \sum_{e \in \text{arr}(\alpha), n(e) \neq 0} \text{sign}(n(e)) t^{|n(e)|}$$

where  $\text{sign}(n) = 1$  for positive  $n \in \mathbb{Z}$  and  $\text{sign}(n) = -1$  for negative  $n \in \mathbb{Z}$ . Therefore

$$\sum_{k \geq 1} k u_k(\alpha) t^{k-1} = u'(\alpha) = \sum_{e \in \text{arr}(\alpha), n(e) \neq 0} n(e) t^{|n(e)|-1} = \sum_{e \in \text{arr}(\alpha)} n(e) t^{|n(e)|-1}.$$

Substituting  $t = 1$ , we obtain

$$\sum_{k \geq 1} k u_k(\alpha) = u'(1) = \sum_{e \in \text{arr}(\alpha)} n(e) = 0.$$

The last equality follows from the fact that if an arrow  $f$  links an arrow  $e$  positively, then  $e$  links  $f$  negatively.

**3.3. Examples.** 1. For positive integers  $p, q$ , we define  $\alpha_{p,q}$  to be the lattice-looking virtual string formed by a Euclidean circle in  $\mathbb{R}^2$  with counterclockwise orientation,  $p$  disjoint vertical arrows  $e_1, \dots, e_p$  directed upward and numerated from left to right, and  $q$  disjoint horizontal arrows  $e_{p+1}, \dots, e_{p+q}$  crossing  $e_1, \dots, e_p$  from right to left and numerated from bottom to top. (Here we identify arrows with geometric vectors in  $\mathbb{R}^2$  connecting two points of the core circle; the numeration of the arrows is compatible with the counterclockwise order of their tails.) Clearly,  $n(e_i) = q$  for  $i = 1, \dots, p$  and  $n(e_{p+j}) = -p$  for  $j = 1, \dots, q$ . Hence  $u(\alpha_{p,q}) = pt^q - qt^p$ . We conclude that the strings  $\{\alpha_{p,q}\}_{p \neq q}$  are pairwise non-homotopic and homotopically non-trivial. The string  $\alpha_{1,1}$  is homotopically trivial: it is obtained from a trivial string by (b)<sub>s</sub>. For  $p \geq 2$ , we have  $u(\alpha_{p,p}) = 0$ . However  $\alpha_{p,p}$  is homotopically non-trivial as will be shown below.

It follows from the definitions that  $\overline{\alpha_{p,q}} = \alpha_{p,q}$  and  $(\alpha_{p,q})^- = \alpha_{q,p}$ . Thus the string  $\alpha = \alpha_{p,q}$  with  $p \neq q$  is not homotopic to  $\alpha^-, \overline{\alpha^-}$ .

Formula 3.2.1 implies that the strings  $\alpha_{p,1}$  and  $\alpha_{1,p}$  with  $p \geq 2$  have minimal rank in their homotopy classes. We shall prove below that the same holds for all  $\alpha_{p,q}$  with  $(p, q) \neq (1, 1)$ .

2. A permutation  $\sigma$  of the set  $\{1, 2, \dots, m\}$  gives rise to a virtual string  $\alpha_\sigma$  of rank  $m$  as follows. Let  $S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$  be the unit circle with counterclockwise orientation. For  $i = 1, \dots, m$ , let  $a_i$  (resp.  $b_i$ ) be the point of  $S^1$  with real part  $(i-1)/m$  and negative (resp. positive) imaginary part. Then  $\alpha_\sigma$  is formed by  $S^1$  and the  $m$  arrows  $\{(a_i, b_{\sigma(i)})\}_{i=1}^m$ . For the  $i$ -th arrow  $e_i = (a_i, b_{\sigma(i)})$ ,

$$(3.3.1) \quad n(e_i) = \#\{j = i+1, \dots, m \mid \sigma(j) < \sigma(i)\} - \#\{j = 1, \dots, i-1 \mid \sigma(j) > \sigma(i)\}.$$

This allows us to compute the polynomial  $u(\alpha_\sigma)$  directly from  $\sigma$ . This example generalizes the previous one since  $\alpha_{p,q} = \alpha_\sigma$  for the permutation  $\sigma$  of the set  $\{1, 2, \dots, p+q\}$  given by

$$\sigma(i) = \begin{cases} i+q, & \text{if } 1 \leq i \leq p \\ i-p, & \text{if } p < i \leq p+q. \end{cases}$$

**3.4. Properties of  $u$ .** We point out a few simple properties of the polynomial  $u$ . For a virtual string  $\alpha$ , we have  $u(\alpha) = u(\overline{\alpha})$ . This follows from the fact that if two arrows are linked positively (resp. negatively), then the reversed arrows are also linked positively (resp. negatively). The transformation  $\alpha \mapsto \alpha^-$  transforms positively linked pairs of arrows into negatively linked pairs and vice versa. Therefore  $u(\alpha^-) = -u(\alpha)$ . As an application, we observe that if  $u(\alpha) \neq 0$ , then  $\alpha$  is not homotopic to  $\alpha^-, \overline{\alpha^-}$ .

It is obvious that if a string  $\alpha$  is a product of strings  $\alpha_1$  and  $\alpha_2$ , then  $u(\alpha) = u(\alpha_1) + u(\alpha_2)$ .

**Theorem 3.4.1.** *An integral polynomial  $u(t)$  can be realized as the  $u$ -polynomial of a virtual string if and only if  $u(0) = u'(1) = 0$ .*

*Proof.* We need only to prove the sufficiency of the condition  $u(0) = u'(1) = 0$ . The proof goes by induction on the degree of  $u$ . If this degree is  $\leq 1$ , then  $u = 0$  is realized by a trivial virtual string. Assume that our claim is true for polynomials of degree  $< m$  where  $m \geq 2$ . Let  $u(t)$  be a polynomial of degree  $m$  with highest term  $at^m$  where  $a \in \mathbb{Z}$  and  $a \neq 0$ . Then  $v(t) = u(t) - a(t^m - mt)$  is a polynomial of degree  $< m$  with  $v(0) = v'(1) = 0$ . By the inductive assumption,  $v(t)$  is realizable as the  $u$ -polynomial of a string. By Example 3.3, the polynomial  $t^m - mt$  is also realizable. Taking a product of strings we observe that the sum of realizable polynomials is realizable. Hence for  $a > 0$ , the polynomial  $u(t) = v(t) + a(t^m - mt)$  is realizable. If  $a < 0$ , then this argument shows that  $-u(t)$  is realizable by a string,  $\alpha$ . Then  $u(t)$  is realized by  $\alpha^-$ .  $\square$

**3.5. Computation for curves.** We compute the polynomial  $u$  for the string  $\alpha = \alpha_\omega$  underlying a closed curve  $\omega : S \rightarrow \Sigma$  on a surface  $\Sigma$ . The computation goes in terms of the homological intersection form  $B : H_1(\Sigma) \times H_1(\Sigma) \rightarrow \mathbb{Z}$  determined by the orientation of  $\Sigma$ . Here and below  $H_1(\Sigma) = H_1(\Sigma; \mathbb{Z})$ .

Let  $e = (a, b)$  be an arrow of  $\alpha$ . Then  $\omega(a) = \omega(b)$  so that  $\omega$  transforms the arcs  $ab, ba \subset S$  into loops  $\omega(ab), \omega(ba)$  in  $\Sigma$ . Set  $[e] = [\omega(ab)] \in H_1(\Sigma)$  and  $[e]^* = [\omega(ba)] \in H_1(\Sigma)$  where the square brackets on the right-hand side stand for the homology class of a loop. We compute the intersection number  $B([e], [e]^*) \in \mathbb{Z}$ . The loops  $\omega(ab), \omega(ba)$  intersect transversely except at their common origin  $\omega(a) = \omega(b)$ . Drawing a picture of  $\omega(ab), \omega(ba)$  in a neighborhood of  $\omega(a) = \omega(b)$ , one observes that a small deformation makes these loops disjoint in this neighborhood. The transversal intersections of  $\omega(ab), \omega(ba)$  bijectively correspond to the

arrows of  $\alpha$  linked with  $e$ , i.e., the arrows connecting an interior point of  $ab$  with an interior point of  $ba$ . The intersection sign at such an intersection is  $+1$  if the tail of the corresponding arrow lies on  $ab$  and is  $-1$  otherwise. Adding these signs, we obtain that  $B([e], [e]^*) = n(e)$ . This formula can be rewritten in a more convenient form. Set  $s = [\omega] = [\omega(S)] \in H_1(\Sigma)$ . Observe that  $s = [e] + [e]^*$  and therefore

$$B([e], [e]^*) = B([e], s - [e]) = B([e], s) - B([e], [e]) = B([e], s).$$

Thus

$$(3.5.1) \quad n(e) = B([e], s).$$

Therefore for any  $k \geq 1$ ,

$$u_k(\alpha) = \#\{e \in \text{arr}(\alpha) \mid B([e], s) = k\} - \#\{e \in \text{arr}(\alpha) \mid B([e], s) = -k\}$$

and

$$u(\alpha) = \sum_{e \in \text{arr}(\alpha), B([e], s) \neq 0} \text{sign}(B([e], s)) t^{|B([e], s)|}.$$

Using the bijective correspondence between the set  $\text{arr}(\alpha)$  and the set  $\mathfrak{N}(\omega)$  of the double points of  $\omega$ , we can rewrite the previous formula as

$$(3.5.2) \quad u(\alpha) = \sum_{x \in \mathfrak{N}(\omega), B([\omega_x], s) \neq 0} \text{sign}(B([\omega_x], s)) t^{|B([\omega_x], s)|}$$

where for  $x \in \mathfrak{N}(\omega)$ , we let  $\omega_x : [0, 1] \rightarrow \Sigma$  be the loop beginning at  $x$  and following along  $\omega$  until the first return to  $x$  and such that the pair (a positive tangent vector of  $\omega_x$  at 0, a positive tangent vector of  $\omega_x$  at 1) is a positive basis in the tangent space of  $x$ .

**Theorem 3.5.1.** *The polynomial  $u$  of a slice virtual string is equal to 0.*

*Proof.* Let  $\alpha_0$  be a slice virtual string realized by a closed curve  $\omega_0$  on the boundary of a handlebody  $H$  such that  $\omega_0$  is contractible in  $H$ . By a *meridian* of  $H$  we mean an embedding  $S^1 \hookrightarrow \partial H$  which extends to an embedding of a 2-disc into  $H$ . Pick a base point  $*$  in  $\partial H$  and let  $1 \in S^1$  be the base point of  $S^1$ . The kernel of the inclusion homomorphism  $\pi_1(\partial H, *) \rightarrow \pi_1(H, *)$  is normally generated by the homotopy classes of meridians of  $H$ . Therefore  $\omega_0$  is homotopic to a loop  $\prod_{i=1}^n r_i m_i r_i^{-1}$  where  $m_1, \dots, m_n : S^1 \hookrightarrow \partial H$  are meridians of  $H$  and  $r_i : [0, 1] \rightarrow \partial H$  is a path leading from  $*$  to  $m_i(1)$ . Deforming slightly these meridians and paths we can assume that the images of  $m_1, \dots, m_n$  are disjoint simple closed curves, that  $r_1, \dots, r_n$  meet these curves and each other transversely, and that in a neighborhood of  $*$  the paths  $r_1, \dots, r_n$  look like radii going out of  $*$  in the cyclic order  $r_1, \dots, r_n$ . Pushing  $r_i$  slightly to the left (resp. to the right) we obtain a “parallel” path  $r_i^+$  (resp.  $r_i^-$ ). Doing it carefully we can assume that  $r_i^+(0) = r_{i-1}^-(0)$  for  $i = 1, \dots, n$  so that in a neighborhood of  $*$  the paths  $r_1^+, r_1^-, \dots, r_n^+, r_n^-$  form  $n$  disjoint embedded arcs approximating the  $n$  radii above from both sides. We also assume that  $r_i^+(1)$  (resp.  $r_i^-(1)$ ) is a point of  $m_i(S^1)$  lying just after (resp. just before)  $m_i(1)$ . Let  $m'_i : [0, 1] \rightarrow \partial H$  be an arc leading from  $r_i^+(1)$  to  $r_i^-(1)$  along  $m_i(S^1)$ . Then  $\omega = r_1^+ m'_1 (r_1^-)^{-1} \cdots r_n^+ m'_n (r_n^-)^{-1}$  is a generic loop in  $\partial H$  homotopic to  $\omega_0$ . Let  $\alpha$  be the underlying virtual string of  $\omega$ . The loop  $\omega$  has two types of self-crossings: each crossing of  $r_i$  with  $m_j$  (possibly  $i = j$ ) gives rise to two self-crossings  $x, y$  of  $\omega$ ; each crossing of  $r_i$  with  $r_j$  gives rise to four self-crossings  $x, y, z, t$  of  $\omega$ . We claim that in the first case  $x, y$  contribute opposite terms to the right-hand side of Formula 3.5.2; in the second case the points  $x, y, z, t$  can be partitioned into two pairs each contributing opposite terms to the right-hand side of 3.5.2. This would imply that  $u(\alpha) = 0$  and since  $\alpha$  is homotopic to  $\alpha_0$ , we also have  $u(\alpha_0) = 0$ . To prove our claim in the first case it suffices to check that  $B([\omega_x], s) = -B([\omega_y], s)$  where  $s = [\omega] \in H_1(\partial H)$ . It follows from the definitions of  $\omega_x, \omega_y$  that  $[\omega_x] + [\omega_y] = s \pm [m_i]$  for  $i \neq j$  and  $[\omega_x] + [\omega_y] = s \pm (s - [m_i])$  for  $i = j$ . The sign  $\pm$  in these formulas depends on the intersection sign at the crossing of  $r_i$  with  $m_j$ . Note that  $B(s, s) = B([m_i], s) = 0$  since  $s = [m_1] + \cdots + [m_n]$  and the meridians  $m_1, \dots, m_n$  are disjoint. Hence  $B([\omega_x], s) + B([\omega_y], s) = 0$ . In the second case we can numerate the points  $x, y, z, t$  so that  $[\omega_x] + [\omega_y] = s + [m_i]$ ,  $[\omega_z] + [\omega_t] = s - [m_i]$  and apply the same argument as above.  $\square$

**Corollary 3.5.2.** *For any  $p \neq q$ , the string  $\alpha_{p,q}$  is not slice.*

**3.6. Remarks.** 1. It is easy to verify that all virtual strings of rank 3 are either homotopically trivial or homeomorphic to  $\alpha_{1,2}, \alpha_{2,1}$ . Note that the homotopy classes of  $\alpha_{1,2}, \alpha_{2,1}$  are distinguished already by  $u_1$ ; indeed  $u_1(\alpha_{1,2}) = -2$  and  $u_1(\alpha_{2,1}) = 2$ .

2. For an open string  $\mu$  with core manifold  $[0, 1]$ , we can define two polynomials  $u^+(\mu)$  and  $u^-(\mu)$ . Observe that the set  $\text{arr}(\mu)$  of arrows of  $\mu$  is a disjoint union  $\text{arr}^+(\mu) \cup \text{arr}^-(\mu)$  where  $\text{arr}^+(\mu)$  (resp.  $\text{arr}^-(\mu)$ ) is the set of arrows  $(a, b) \in \text{arr}(\mu)$  with  $a, b \in [0, 1]$  such that  $a < b$  (resp.  $b < a$ ). For  $e \in \text{arr}(\mu)$ , set  $n(e) = n(e^{cl}) \in \mathbb{Z}$  where  $e^{cl}$  is the corresponding arrow of the closure,  $\mu^{cl}$ , of  $\mu$ . For  $k \geq 1$ , set

$$u_k^\pm(\mu) = \#\{e \in \text{arr}^\pm(\mu) \mid n(e) = k\} - \#\{e \in \text{arr}^\mp(\mu) \mid n(e) = -k\} \in \mathbb{Z}.$$

This number and the polynomials  $u^\pm(\mu) = \sum_{k \geq 1} u_k^\pm(\mu) t^k$  are homotopy invariants of  $\mu$ . Clearly,  $u(\mu^{cl}) = u^+(\mu) + u^-(\mu)$ . Using  $u^\pm$ , it is easy to give examples of non-homotopic open strings with homotopic closures.

3. For a virtual string  $\alpha$  and a positive integer  $p$ , we define a virtual string  $\alpha(p)$  as follows. Let us identify the core circle of  $\alpha$  with  $S^1 \subset \mathbb{C}$ . Each arrow of  $\alpha$  can be graphically presented by a vector in  $\mathbb{C}$  connecting two points of  $S^1$ . These vectors are mutually transverse. Now, replace each of these vectors, say  $e$ , by  $p$  disjoint parallel vectors  $e_1, \dots, e_p$  running closely to  $e$  and having endpoints on  $S^1$ . This gives a virtual string  $\alpha(p)$  of rank  $pm$  where  $m$  is the rank of  $\alpha$ . It is obvious that  $n(e_1) = \dots = n(e_p) = pn(e)$ . Therefore  $u(\alpha(p))(t) = pu(\alpha)(t^p)$ . In particular, if  $u(\alpha) \neq 0$ , then  $\alpha(p)$  is homotopically non-trivial.

#### 4. GEOMETRIC REALIZATION OF VIRTUAL STRINGS

**4.1. Realization of strings.** We explain here that every virtual string admits a canonical realization by a closed curve on a surface and moreover describe all its realizations.

Let  $\alpha$  be a virtual string of rank  $m$  with core circle  $S$ . Identifying the tail with the head for all arrows of  $\alpha$ , we transform  $S$  into a 1-dimensional CW-complex  $\Gamma = \Gamma_\alpha$ . We thicken  $\Gamma$  to a surface  $\Sigma_\alpha$  as follows. If  $m = 0$ , then  $\Gamma = S$  and we set  $\Sigma_\alpha = S \times [-1, 1]$ . Assume that  $m \geq 1$ . The 0-cells (vertices) of  $\Gamma$  have valency 4 and their number is equal to  $m$ . A neighborhood of a vertex  $v \in \Gamma$  embeds into the unit 2-disc  $D^2 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$  as follows. Suppose that  $v$  is obtained from an arrow  $(a, b)$  where  $a, b \in S$ . Note that any point  $x \in S$  splits its small neighborhood in  $S$  into two oriented arcs, one of them being incoming and the other one being outgoing with respect to  $x$ . Therefore a neighborhood of  $v$  in  $\Gamma$  consists of four arcs which can be identified with small incoming and outgoing arcs of  $a, b$  on  $S$ . We embed this neighborhood into  $D^2$  so that  $v$  goes to the origin and the incoming (resp. outgoing) arcs of  $a, b$  go to the intervals  $[-1, 0] \times 0, 0 \times [-1, 0]$  (resp.  $[0, 1] \times 0, 0 \times [0, 1]$ ), respectively. In this way all vertices of  $\Gamma$  are thickened to copies of  $D^2$  endowed with counterclockwise orientation. Each 1-cell of  $\Gamma$  connects two (possibly coinciding) vertices and is thickened to a ribbon connecting the corresponding 2-discs. The thickening is uniquely determined by the condition that the orientation of these 2-discs extends to their union with the ribbon. Thickening in this way all the vertices and 1-cells of  $\Gamma$  we embed  $\Gamma$  into a surface  $\Sigma_\alpha$ . By construction,  $\Sigma_\alpha$  is a compact connected oriented surface with non-void boundary and Euler characteristic  $\chi(\Sigma_\alpha) = \chi(\Gamma) = -m$ . Composing the natural projection  $S \rightarrow \Gamma$  with the inclusion  $\Gamma \hookrightarrow \Sigma_\alpha$  we obtain a closed curve  $\omega_\alpha : S \rightarrow \Sigma_\alpha$  realizing  $\alpha$ . The construction of  $\Sigma_\alpha$  is well known, see [Fr], [Ca], [CW].

It is clear that for any surface  $\Sigma$  and any (generic) closed curve  $\omega : S \rightarrow \Sigma$  realizing  $\alpha$ , a regular neighborhood of  $\omega(S)$  in  $\Sigma$  is homeomorphic to  $\Sigma_\alpha$ . Moreover, the homeomorphism can be chosen to transform  $\omega$  into  $\omega_\alpha$ . In other words,  $\omega$  can be obtained as a composition of  $\omega_\alpha$  with an orientation-preserving embedding  $\Sigma_\alpha \hookrightarrow \Sigma$ . In particular,  $\Sigma_\alpha$  is a surface of minimal genus containing a closed curve realizing  $\alpha$ . Therefore the genus  $g(\alpha)$  of  $\alpha$  defined in Section 2.5 is equal to the genus of  $\Sigma_\alpha$ . It will be explicitly computed in the next subsection. Note finally that a closed surface of minimal genus containing a curve realizing  $\alpha$  is obtained from  $\Sigma_\alpha$  by gluing 2-discs to all components of  $\partial\Sigma_\alpha$ .

**4.2. Homological computations.** Consider again a virtual string  $\alpha$  of rank  $m$  with core circle  $S$ . Let  $\Gamma = \Gamma_\alpha$ ,  $\Sigma = \Sigma_\alpha$ , and  $\omega = \omega_\alpha : S \rightarrow \Sigma_\alpha$  be the graph, the surface, and the closed curve constructed in the previous subsection. The orientation of  $\Sigma$  determines a homological intersection pairing  $B = B_\alpha : H_1(\Sigma) \times H_1(\Sigma) \rightarrow \mathbb{Z}$ . This bilinear pairing is skew-symmetric and its rank is equal to twice the genus of  $\Sigma$ . Thus

$$g(\alpha) = (1/2) \text{rank} B_\alpha.$$

In particular,  $\alpha$  can be realized by a closed curve on  $S^2$  or  $\mathbb{R}^2$  if and only if  $B_\alpha = 0$ .

Since  $\Gamma$  is a deformation retract of  $\Sigma$ , the inclusion homomorphism  $H_1(\Gamma) \rightarrow H_1(\Sigma)$  is an isomorphism. Since  $\Gamma$  is a connected graph with  $\chi(\Gamma) = -m$ , the group  $H_1(\Gamma) = H_1(\Sigma)$  is a free abelian group of rank



$m + 1$ . We describe a canonical basis in  $H_1(\Sigma)$ . Set  $s = [\omega] \in H_1(\Sigma)$ . For an arrow  $e = (a, b) \in \text{arr}(\alpha)$ , the map  $\omega$  transforms the arc  $ab \subset S$ , leading from  $a$  to  $b$  in the positive direction, into a loop  $\omega(ab)$  in  $\Sigma$ . Set  $[e] = [\omega(ab)] \in H_1(\Sigma)$ . An easy induction on  $m$  shows that  $s \cup \{[e]\}_{e \in \text{arr}(\alpha)}$  is a basis of  $H_1(\Sigma)$ . Our next aim is to compute the matrix of  $B$  in this basis. Note for the record that  $B(x, y) = -B(y, x)$  and  $B(x, x) = 0$  for all elements  $x, y$  of this basis.

By Formula 3.5.1,  $B([e], s) = n(e)$  for any  $e \in \text{arr}(\alpha)$ . To compute the other values of  $B$ , we need more notation. Let  $a, b$  be distinct point of  $S$ . The *interior* of the arc  $ab \subset S$  is the set  $(ab)^\circ = ab - \{a, b\}$ . For any arcs  $ab, cd \subset S$ , we define  $ab \cdot cd \in \mathbb{Z}$  to be the number of arrows of  $\alpha$  with tail in  $(ab)^\circ$  and head in  $(cd)^\circ$  minus the number of arrows of  $\alpha$  with tail in  $(cd)^\circ$  and head in  $(ab)^\circ$ . Note that the arrows with both endpoints in  $(ab)^\circ \cap (cd)^\circ$  appear in this expression twice with opposite signs and therefore cancel out. Clearly,  $ab \cdot cd = -cd \cdot ab$ . In particular,  $ab \cdot ab = 0$ . If  $e = (a, b)$  is an arrow of  $\alpha$ , then it follows from the definitions that  $n(e) = ab \cdot ba$ . More generally, for any arrow  $f = (c, d)$  of  $\alpha$  unlinked with  $e = (a, b)$ ,

$$(4.2.1) \quad n(e) = ab \cdot cd + ab \cdot dc.$$

Note that the arrows  $e = (a, b)$ ,  $f = (c, d)$  never contribute to  $ab \cdot cd$  and to  $ab \cdot dc$  because neither endpoint of  $e$  lies in  $(ab)^\circ$  and neither endpoint of  $f$  lies in  $(cd)^\circ$  or in  $(dc)^\circ$ . Applying Formula 4.2.1 to the string obtained from  $\alpha$  by reversing the arrow  $e$ , we obtain that

$$(4.2.2) \quad n(e) = -ba \cdot cd - ba \cdot dc.$$

**Lemma 4.2.1.** *Let  $e = (a, b)$  and  $f = (c, d)$  be two arrows of  $\alpha$ . Then  $B([e], [f]) = ab \cdot cd + \varepsilon$  where  $\varepsilon = 0$  if  $e$  and  $f$  are unlinked,  $\varepsilon = 1$  if  $f$  links  $e$  positively, and  $\varepsilon = -1$  if  $f$  links  $e$  negatively.*

*Proof.* If  $e = f$ , then  $a = c, b = d$  and all terms of the stated equality are equal to 0. (Note that an arrow is unlinked with itself.) Assume from now on that  $e \neq f$  so that  $a, b, c, d$  are pairwise distinct points of  $S$ .

Suppose first that  $e$  and  $f$  are unlinked. There are four cases to consider depending on whether the endpoints of  $e, f$  lie on  $S$  in the cyclic order (i)  $a, b, c, d$ , or (ii)  $a, b, d, c$ , or (iii)  $a, c, d, b$ , or (iv)  $a, d, c, b$ .

In the case (i), the arcs  $ab, cd \subset S$  are disjoint so that  $[e], [f] \in H_1(\Sigma)$  are represented by transversal loops  $\omega(ab), \omega(cd)$ , respectively. Then  $B([e], [f]) = ab \cdot cd$ , cf. Section 3.5.

In the case (ii), the arcs  $ab, dc \subset S$  are disjoint so that  $[e], [f]^* = s - [f] \in H_1(\Sigma)$  are represented by transversal loops  $\omega(ab), \omega(dc)$ , respectively. Hence  $B([e], [f]^*) = ab \cdot dc$  and

$$B([e], [f]) = B([e], s - [f]^*) = B([e], s) - B([e], [f]^*) = n(e) - ab \cdot dc = ab \cdot cd.$$

In the case (iii), we have  $B([e], [f]) = -B([f], [e]) = -cd \cdot ab = ab \cdot cd$  since the pair  $(f, e)$  satisfies the conditions of (ii).

In the case (iv), the arcs  $ba, dc \subset S$  are disjoint so that  $[e]^* = s - [e], [f]^* = s - [f] \in H_1(\Sigma)$  are represented by transversal loops  $\omega(ba), \omega(dc)$ , respectively. Therefore

$$B([e], [f]) = B([e], s) + B(s, [f]) + B(s - [e], s - [f]) = n(e) - n(f) + B([e]^*, [f]^*) = n(e) - n(f) + ba \cdot dc.$$

It remains to observe that

$$ba \cdot dc = -n(e) - ba \cdot cd = -n(e) + cd \cdot ba = -n(e) + n(f) - cd \cdot ab = -n(e) + n(f) + ab \cdot cd.$$

Suppose that  $f$  links  $e$  positively. Then their endpoints lie on  $S$  in the cyclic order  $a, c, b, d$ . The loops  $X = \omega(ab), Y = \omega(cd)$  representing  $[e], [f] \in H_1(\Sigma)$  are not transversal since both contain  $\omega(cb)$ . Pushing  $Y$  slightly to its left in  $\Sigma$ , we obtain a loop,  $Y^+$ , transversal to  $X$ . It is understood that the point  $\omega(c) = \omega(d) \in Y$  is pushed to a point lying between  $\omega(ac)$  and  $\omega(da)$  in a small neighborhood of  $\omega(c) = \omega(d)$ . Introducing coordinates  $(x, y)$  in this neighborhood we can locally identify  $X, Y, Y^+$  with the axis  $y = 0$ , the union of two half-lines  $x = 0, y \leq 0$  and  $y = 0, x \geq 0$ , and the union of two half-lines  $x = -1, y \leq 1$  and  $y = 1, x \geq -1$ , respectively. To compute the intersection number  $B([e], [f]) = X \cdot Y = X \cdot Y^+$ , we split the set  $X \cap Y^+$  into four disjoint subsets. The first of them consists of a single point near  $\omega(c) = \omega(d)$ , given in the coordinates above by  $x = -1, y = 0$ . This point contributes 1 to  $X \cdot Y^+$ . The second subset of  $X \cap Y^+$  is  $\omega(ac) \cap Y^+$ ; its points are numerated by arrows of  $\alpha$  with one endpoint in the interior of  $ac$  and the other endpoint in the interior of  $cd$ . The contribution of these crossings to  $X \cdot Y^+$  is equal to  $ac \cdot cd$ . The third subset of  $X \cap Y^+$  is numerated by the crossings of  $\omega(cb)$  with the part of  $Y^+$  obtained by pushing  $\omega(bd) \subset Y$  to the left; they are numerated by arrows of  $\alpha$  with one endpoint in the interior of  $cb$  and the other endpoint in the interior of  $bd$ . The contribution of these crossings to  $X \cdot Y^+$  is  $cb \cdot bd$ . The forth subset of  $X \cap Y^+$  is numerated by

the self-crossings of  $\omega(cb)$ : each of them gives rise to two points of  $X \cap Y^+$  with opposite intersection signs. Therefore this forth subset contributes 0 to  $X \cdot Y^+$ . Summing up these contributions we obtain

$$\begin{aligned} B([e], [f]) &= 1 + ac \cdot cd + cb \cdot bd + 0 = ac \cdot cd + cb \cdot cb + cb \cdot bd + 1 \\ &= ac \cdot cd + cb \cdot cd + 1 = ab \cdot cd + 1. \end{aligned}$$

If  $f$  links  $e$  negatively, then  $e$  links  $f$  positively and by the results above,

$$B([e], [f]) = -B([f], [e]) = -(cd \cdot ab + 1) = ab \cdot cd - 1.$$

□

**4.3. Examples.** (1) Consider the string  $\alpha = \alpha_{p,q}$  with  $p, q \geq 1$  introduced in Section 3.3.1. Recall the arrows  $e_1, \dots, e_{p+q}$  of  $\alpha$ . We compute the matrix of the bilinear form  $B = B_\alpha : H_1(\Sigma_\alpha) \times H_1(\Sigma_\alpha) \rightarrow \mathbb{Z}$  with respect to the basis  $s \cup \{[e_i]\}_{i=1}^{p+q}$ . By Formula 3.5.1,  $B([e_i], s) = q$  for  $i = 1, \dots, p$  and  $B([e_{p+j}], s) = -p$  for  $j = 1, \dots, q$ . Each pair of arrows  $e_i, e_{i'}$  with  $i, i' = 1, \dots, p$  is unlinked and by Lemma 4.2.1,  $B([e_i], [e_{i'}]) = 0$ . Similarly, each pair of arrows  $e_{p+j}, e_{p+j'}$  with  $j, j' = 1, \dots, q$  is unlinked and  $B([e_{p+j}], [e_{p+j'}]) = 0$ . The arrow  $e_{p+j}$  links  $e_i$  positively and by Lemma 4.2.1,  $B([e_i], [e_{p+j}]) = (p-i) + (q-j) + 1$ . It is easy to compute that the rank of  $B$  is equal to 2 if  $p = q = 1$ , to 6 if  $\min(p, q) \geq 3$ , and to 4 in all the other cases. The genus  $g(\alpha)$ , as we know, is half of this rank. In particular,  $g(\alpha_{1,1}) = 1$  which shows that the genus is not a homotopy invariant.

(2) Consider the string  $\alpha = \alpha_\sigma$  defined in Section 3.3.2 for a permutation  $\sigma$  of the set  $\{1, 2, \dots, m\}$ . Recall the arrows  $e_1, \dots, e_m$  of  $\alpha$ . We compute the matrix of the bilinear form  $B = B_\alpha : H_1(\Sigma_\alpha) \times H_1(\Sigma_\alpha) \rightarrow \mathbb{Z}$  with respect to the basis  $s \cup \{[e_i]\}_{i=1}^m$ . The number  $B([e_i], s) = n(e_i)$  is computed by Formula 3.3.1. Pick two indices  $i, j$  with  $1 \leq i < j \leq m$ . Lemma 4.2.1 implies that if  $\sigma(i) < \sigma(j)$ , then

$$B([e_i], [e_j]) = \#\{k \mid i < k < j, \sigma(j) < \sigma(k)\} - \#\{k \mid j < k \leq m, \sigma(i) < \sigma(k) < \sigma(j)\}.$$

If  $\sigma(j) < \sigma(i)$ , then

$$B([e_i], [e_j]) = \#\{k \mid i < k < j, \sigma(j) < \sigma(k)\} + \#\{k \mid j < k \leq m, \sigma(j) < \sigma(k) < \sigma(i)\} + 1.$$

**4.4. Homotopy of strings re-examined.** The fact that all strings can be realized by curves on surfaces allows us to reformulate the notion of homotopy of strings entirely in terms of homotopies of curves. For strings  $\alpha, \beta$ , we write  $\alpha \sim \beta$  if these two strings can be realized by homotopic closed curves on the same surface. The relation  $\sim$  is reflexive and symmetric but not transitive. The next lemma shows that the relation of homotopy is precisely the equivalence relation generated by  $\sim$ .

**Lemma 4.4.1.** *Two strings  $\alpha, \beta$  are homotopic if and only if there is a sequence of strings  $\alpha_1 = \alpha, \alpha_2, \dots, \alpha_n = \beta$  such that  $\alpha_i \sim \alpha_{i+1}$  for  $i = 1, \dots, n-1$ .*

*Proof.* As we know, the underlying virtual strings of homotopic closed curves on a surface are themselves homotopic. Therefore if there is a sequence of strings  $\alpha_1 = \alpha, \alpha_2, \dots, \alpha_n = \beta$  such that  $\alpha_i \sim \alpha_{i+1}$  for  $i = 1, \dots, n-1$ , then  $\alpha$  is homotopic to  $\beta$ . To prove the converse, it suffices to show that if  $\beta$  is obtained from  $\alpha$  by a homotopy move (a)<sub>s</sub>, (b)<sub>s</sub>, or (c)<sub>s</sub>, then  $\alpha \sim \beta$ .

Let  $S$  be the core circle of  $\alpha$  and  $\omega : S \rightarrow \Sigma$  be a curve realizing  $\alpha$  on a surface  $\Sigma$ . Pick distinct points  $a, b \in S$  such that the arc  $ab \subset S$  does not contain endpoints of  $\alpha$ . Let  $\beta$  be obtained from  $\alpha$  by the move (a)<sub>s</sub> adding the arrow  $(a, b)$ . Attaching to  $\omega$  a small curl on the right of the arc  $\omega(ab)$ , we obtain a closed curve  $\omega' : S \rightarrow \Sigma$  realizing  $\beta$ . Clearly,  $\omega'$  is homotopic to  $\omega$  in  $\Sigma$ . Hence  $\alpha \sim \beta$ .

Pick two arcs  $x, y$  on  $S$  disjoint from each other and from the endpoints of  $\alpha$ . Let  $a, a'$  be the endpoints of  $x$  (in an arbitrary order) and  $b, b'$  be the endpoints of  $y$ . Let  $\beta$  be obtained from  $\alpha$  by the move (b)<sub>s</sub> adding to  $\alpha$  the arrows  $(a, b)$  and  $(b', a')$ . Let  $D_x, D_y \subset \Sigma - \omega(S)$  be two small closed discs lying near the arcs  $\omega(x), \omega(y)$ , respectively. Removing the interiors of these discs from  $\Sigma$  and gluing the circles  $\partial D_x, \partial D_y$  along an orientation-reversing homeomorphism, we obtain a new (oriented) surface,  $\Sigma'$ , containing  $\omega(S)$ . In  $\Sigma'$  the arcs  $\omega(x)$  and  $\omega(y)$  are adjacent to the component of  $\Sigma' - \omega(S)$  containing  $\partial D_x = \partial D_y$ . We can push  $\omega(x)$  across this component towards  $\omega(y)$  and eventually across  $\omega(y)$ . This gives a curve  $\omega' : S \rightarrow \Sigma'$  realizing  $\beta$  and homotopic to  $\omega : S \rightarrow \Sigma'$ . Hence  $\alpha \sim \beta$ . Note that the four possible forms of the move (b)<sub>s</sub> (depending on whether  $x$  leads from  $a$  to  $a'$  or from  $a'$  to  $a$  and similarly for  $y$ ) are realized by choosing  $D_x, D_y$  on the left or on the right of  $\omega(x), \omega(y)$ .

Suppose that  $\alpha$  has three arrows  $(a^+, b), (b^+, c), (c^+, a)$  where  $a, a^+, b, b^+, c, c^+ \in S$  such that the (positively oriented) arcs  $aa^+, bb^+, cc^+$  are disjoint from each other and from the other endpoints of  $\alpha$ . Let  $\beta$  be obtained

from  $\alpha$  by the move  $(c)_s$  replacing the arrows  $(a^+, b), (b^+, c), (c^+, a)$  with the arrows  $(a, b^+), (b, c^+), (c, a^+)$ . Consider the canonical realization  $\omega_\alpha : S \rightarrow \Sigma_\alpha$  of  $\alpha$ . Observe that the arcs  $\omega_\alpha(aa^+), \omega_\alpha(bb^+), \omega_\alpha(cc^+)$  form a simple closed curve in  $\Sigma_\alpha$  isotopic to a boundary component of  $\Sigma_\alpha$ . Gluing a 2-disc  $D$  to this boundary component we embed  $\Sigma_\alpha$  into a bigger surface,  $\Sigma$ . Pushing the branch  $\omega_\alpha(aa^+)$  across  $D \subset \Sigma$  and then across the double point  $\omega(b^+) = \omega(c)$ , we obtain a curve  $\omega' : S \rightarrow \Sigma$  realizing  $\beta$  and homotopic to  $\omega_\alpha : S \rightarrow \Sigma$ . Hence  $\alpha \sim \beta$ .  $\square$

**4.5. Adams operations.** We can define “Adams operations”  $\{\psi^n\}_{n \in \mathbb{Z}}$  on the set  $\mathcal{S}$  of homotopy classes of virtual strings. Let  $\alpha$  be a string. Replacing  $\alpha$  by a homeomorphic string, we can identify its core circle with  $S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ . Consider a curve  $\omega : S^1 \rightarrow \Sigma$  realizing  $\alpha$  on a surface  $\Sigma$ . The mapping  $S^1 \rightarrow \Sigma$  sending  $z \in S^1$  to  $\omega(z^n)$  is homotopic to a generic curve  $S^1 \rightarrow \Sigma$ . We define  $\psi^n(\alpha) \in \mathcal{S}$  to be the homotopy class of its underlying string. Lemma 4.4.1 implies that  $\psi^n : \mathcal{S} \rightarrow \mathcal{S}$  is a well defined mapping. Clearly,  $\psi^1(\alpha) = \alpha$ ,  $\psi^{-1}(\alpha) = \alpha^-$ , and  $\psi^{mn} = \psi^m \circ \psi^n$  for any  $m, n \in \mathbb{Z}$ . As an exercise, the reader may check that  $u(\psi^n(\alpha)) = \text{sign}(n) n^2 u(\alpha)$ .

## 5. COMBINATORICS OF CLOSED CURVES ON THE 2-SPHERE

**5.1. Gauss words.** In this section, an *alphabet* is a finite set and *letters* are its elements. A *word* in an alphabet is a finite sequence of letters. The words will be always considered up to circular permutations. A *Gauss word* is a word in an alphabet in which all letters of the alphabet occur exactly twice. Each virtual string gives rise to a Gauss word as follows. Label the arrows of the string with different letters. Traverse the core circle of the string in the positive direction and write down the label of an arrow each time we cross its endpoint. This gives a Gauss word well defined up to circular permutations (and the choice of letters). It is clear that the Gauss word associated in the Introduction with a closed curve on  $\mathbb{R}^2$  (or more generally on any surface) coincides with the Gauss word associated with the underlying virtual string.

Gauss [Ga] studied the following question: when a Gauss word can be realized by a closed curve on the plane? He gave a necessary condition (condition (i) in Theorem 5.3.1 below) which he knew not to be sufficient. Later this question was studied by several authors, see [CW] and M. L. Marx’ review MR2000i:05056 for references. A most elegant solution to this question was given by P. Rosenstiehl [Ro], see also [RR] and Corollary 5.3.2 below.

Note that a Gauss word is realizable by a closed curve on  $\mathbb{R}^2$  if and only if it is realizable by a closed curve on the 2-sphere  $S^2$ . We shall focus on curves on  $S^2$  rather than on  $\mathbb{R}^2$ .

**5.2. Bipartitions.** A *bipartition* (or a *partition into two sets*) of a set  $E$  is a non-ordered pair of disjoint (possibly empty) subsets of  $E$  whose union is  $E$ . We explain now that virtual strings and closed curves on surfaces naturally give rise to bipartitions. Consider a virtual string  $\alpha$  with core circle  $S$ . For arrows  $e = (a, b)$  and  $f = (c, d)$  of  $\alpha$ , we define  $q(e, f) \in \mathbb{Z}$  to be the number of arrowheads of (the arrows of)  $\alpha$  lying on the semi-open arc  $ac - \{a\} \subset S$  minus the number of arrowtails of  $\alpha$  lying on  $ac - \{a\}$ . If  $e = f$ , then by definition  $q(e, f) = 0$ . It is easy to check that

$$q(e, f) + q(f, e) = 0, \quad q(e, f) + q(f, g) + q(g, e) = 0$$

for any  $e, f, g \in \text{arr}(\alpha)$ . We use  $q \pmod{2}$  to define an equivalence relation on  $\text{arr}(\alpha)$ : two arrows  $e, f \in \text{arr}(\alpha)$  are *equivalent* if  $q(e, f) \equiv 0 \pmod{2}$ . This can be reformulated in simpler terms: the arrows  $e = (a, b)$  and  $f = (c, d)$  are equivalent if either  $e = f$  or the number of endpoints of  $\alpha$  lying in the interior of the arc  $ac \subset S$  is odd. It is obvious that this relation on  $\text{arr}(\alpha)$  has at most two equivalence classes. They form a partition of  $\text{arr}(\alpha)$  into two subsets (one of them may be empty). If the arrows of  $\alpha$  are labeled by different letters, then this bipartition of  $\text{arr}(\alpha)$  induces a bipartition of the set of letters. Thus  $\alpha$  gives rise to a pair (a Gauss word, a bipartition of the alphabet).

Applying this construction to the underlying virtual string of a closed curve  $\omega$  on a surface, we obtain a bipartition of the set of double points of  $\omega$ . If the double points of  $\omega$  are labeled by different letters, then this induces a bipartition of the set of letters. Thus  $\omega$  gives rise to a pair (a Gauss word, a bipartition of the alphabet). For curves on  $S^2$ , we shall give a geometric re-formulation of this bipartition in Remark 5.6.1.

**5.3. Curves on  $S^2$ .** It is natural to ask whether the pair (the Gauss word, the bipartition of the alphabet) associated with a curve on a surface is a full homeomorphism invariant of the curve and what values it can take. We shall answer these questions for curves on  $S^2$ .

We begin with definitions. Let  $w$  be a word in an alphabet  $E$ . We say that two (distinct) letters  $i, j \in E$  are  $w$ -interlaced if  $w$  has the form  $i\dots j\dots i\dots j\dots$  up to a circular permutation. For  $i \in E$ , denote by  $w_i$  the set of letters  $w$ -interlaced with  $i$ . A bipartition of  $E$  is *compatible* with  $w$  if it satisfies the following condition: for any  $w$ -interlaced  $i, j \in E$ , we have  $\#(w_i \cap w_j) \equiv 0 \pmod{2}$  if  $i, j$  belong to different subsets of the bipartition and  $\#(w_i \cap w_j) \equiv 1 \pmod{2}$  if  $i, j$  belong to the same subset.

**Theorem 5.3.1.** *A pair (a Gauss word  $w$  in an alphabet  $E$ , a bipartition of  $E$ ) is realizable by a closed curve on  $S^2$  if and only if the following three conditions are satisfied:*

- (i) *for all  $i \in E$ , the set  $w_i$  has an even number of elements:  $\#(w_i) \equiv 0 \pmod{2}$ ;*
- (ii) *if  $i, j \in E$  are not  $w$ -interlaced, then  $\#(w_i \cap w_j) \equiv 0 \pmod{2}$ ;*
- (iii) *the bipartition of  $E$  is compatible with  $w$ .*

*Proof.* We begin with preliminary computations. Let  $\alpha$  be a virtual string of rank  $m \geq 1$  with core circle  $S$ . Suppose that the arrows of  $\alpha$  are labeled by elements of a set  $E$  so that each  $i \in E$  appears as a label of a unique arrow  $e_i = (a_i, b_i)$  of  $\alpha$  where  $a_i, b_i \in S$ . As explained above, this gives a Gauss word  $w = w(\alpha)$  in the alphabet  $E$ . Observe that letters  $i, j \in E$  are  $w$ -interlaced if and only if  $e_i, e_j$  are linked. Therefore for any  $i \in E$ , the number  $\#(w_i)$  is the number of arrows of  $\alpha$  linked with  $e_i$ . Hence

$$\#(w_i) \equiv n(e_i) \pmod{2}.$$

Similarly, for any  $i, j \in E$ , we have  $\#(w_i \cap w_j) \equiv n_{i,j} \pmod{2}$  where  $n_{i,j}$  is the number of arrows of  $\alpha$  linked with both  $e_i$  and  $e_j$ . We now relate  $n_{i,j}$  to the number  $B([e_i], [e_j])$  defined in Section 4.2. We shall assume that  $n(e_i) \equiv n(e_j) \equiv 0 \pmod{2}$  (this will be enough for our aims). Applying Formulas 4.2.1, 4.2.2 to  $e = e_i, e_j$  we obtain that

$$(5.3.1) \quad a_i b_i \cdot a_j b_j \equiv a_i b_i \cdot b_j a_j \equiv b_i a_i \cdot a_j b_j \equiv b_i a_i \cdot b_j a_j \pmod{2}.$$

If  $i, j$  are not  $w$ -interlaced, then one of the arcs  $a_i b_i, b_i a_i$ , say  $\beta$ , is disjoint from one of the arcs  $a_j b_j, b_j a_j$ , say  $\gamma$ . It is obvious that  $n_{i,j} \equiv \beta \cdot \gamma \pmod{2}$ . Lemma 4.2.1 and Formula 5.3.1 imply that

$$(5.3.2) \quad \#(w_i \cap w_j) \equiv n_{i,j} \equiv \beta \cdot \gamma \equiv a_i b_i \cdot a_j b_j = B([e_i], [e_j]) \pmod{2}.$$

Suppose now that  $i, j$  are  $w$ -interlaced. It will be enough for our aims to consider the case where the endpoints of  $e_i, e_j$  lie on  $S$  in the cyclic order  $a_i, a_j, b_i, b_j$ . Then  $n_{i,j} \equiv a_i a_j \cdot b_i b_j + a_j b_i \cdot b_j a_i \pmod{2}$ . Let  $r$  be the number of endpoints of  $\alpha$  lying on the arc  $a_j b_i$ . It is easy to see that  $r \equiv a_j b_i \cdot b_i a_j \pmod{2}$ . The number  $q(e_i, e_j) \in \mathbb{Z}$  defined in Section 5.2 satisfies

$$q(e_i, e_j) + r \equiv \#(w_i) \equiv n(e_i) \equiv 0 \pmod{2}.$$

Therefore  $q(e_i, e_j) \equiv r \equiv a_j b_i \cdot b_i a_j \pmod{2}$ . Using Lemma 4.2.1, we obtain

$$(5.3.3) \quad \begin{aligned} B([e_i], [e_j]) &= a_i b_i \cdot a_j b_j + 1 = a_i a_j \cdot a_j b_i + a_i a_j \cdot b_i b_j + a_j b_i \cdot b_i b_j + 1 \\ &\equiv -a_i a_j \cdot a_j b_i + a_i a_j \cdot b_i b_j + a_j b_i \cdot b_i b_j + 2 a_j b_i \cdot b_j a_i + 1 \\ &= (a_j b_i \cdot a_i a_j + a_j b_i \cdot b_i b_j + a_j b_i \cdot b_j a_i) + (a_i a_j \cdot b_i b_j + a_j b_i \cdot b_j a_i) + 1 \\ &\equiv a_j b_i \cdot b_i a_j + n_{i,j} + 1 \equiv q(e_i, e_j) + \#(w_i \cap w_j) + 1 \pmod{2}. \end{aligned}$$

We can now prove the necessity of the conditions (i) – (iii) of the theorem. Suppose that a Gauss word  $w$  in an alphabet  $E$  is realized by a closed curve  $\omega : S^1 \rightarrow S^2$ . Let  $\alpha$  be the underlying virtual string of  $\omega$ . Thus,  $w = w(\alpha)$  for an appropriate bijective labelling  $E \rightarrow \text{arr}(\alpha), i \mapsto e_i$ . By Formula 3.5.1,  $n(e_i)$  is the intersection number of two cycles lying in a neighborhood of  $\omega(S^1)$  in  $S^2$ . Since the intersection number of any two cycles in  $S^2$  is zero,  $n(e_i) = 0$  and  $\#(w_i) \equiv n(e_i) \equiv 0 \pmod{2}$  for all  $i \in E$ . Condition (ii) follows similarly from Formula 5.3.2. It remains to verify that the bipartition of  $E$  induced by  $\alpha$  is compatible with  $w$ . Let  $i, j \in E$  be  $w$ -interlaced. Permuting if necessary  $i$  and  $j$  we can assume that the endpoints of  $e_i = (a_i, b_i)$ ,  $e_j = (a_j, b_j)$  lie on  $S^1$  in the cyclic order  $a_i, a_j, b_i, b_j$ . Formula 5.3.3 implies that  $q(e_i, e_j) + \#(w_i \cap w_j) + 1 \equiv 0 \pmod{2}$ . Hence  $\#(w_i \cap w_j) \equiv 0 \pmod{2}$  if and only if  $q(e_i, e_j) \equiv 1 \pmod{2}$ , i.e., if and only if  $i, j$  belong to different equivalence classes.

To accomplish the proof of the theorem, we need a general construction of strings from Gauss words and bipartitions. Assume that we have a Gauss word  $w = z_1 z_2 \dots z_{2m}$  in an alphabet  $E$  with  $\#E = m$  and a bipartition  $E = X \cup Y$ . Assume that  $w$  satisfies the “parity condition” (i) of the theorem. We construct a string  $\alpha$  giving rise to this word and this bipartition. The core circle of  $\alpha$  is the circle  $S = \mathbb{R} \cup \{\infty\}$  with orientation extending the right-handed orientation on  $\mathbb{R}$ . The set of arrow endpoints of  $\alpha$  is the set  $\{1, 2, \dots, 2m\} \in \mathbb{R} \subset S$ . This set is partitioned into  $m$  pairs: two points  $a, b \in \{1, 2, \dots, 2m\}$  form a pair if

$z_a = z_b$ . We order each such pair  $\{a, b\}$  as follows. By the parity condition (i), we have  $a - b \equiv 1 \pmod{2}$ . Therefore one of the numbers  $a, b$  is even and the other one is odd. If  $z_a = z_b \in X$ , then we put the odd one on the first place and the even one on the second place. If  $z_a = z_b \in Y$ , then we do the opposite. The string  $\alpha$  is the circle  $S$  with these  $m$  ordered pairs of points. It follows from the definitions that the Gauss word of  $\alpha$  is  $w$  and the induced bipartition of  $E$  is  $E = X \cup Y$ .

To establish the sufficiency of the conditions (i) – (iii) of the theorem, we need only to show that if  $w$  satisfies (i), (ii) and the bipartition  $E = X \cup Y$  is compatible with  $w$ , then the string  $\alpha$  constructed in the previous paragraph is realizable by a closed curve in  $S^2$ . It suffices to show that the surface  $\Sigma = \Sigma_\alpha$  is a disc with holes; then  $\Sigma$  embeds in  $S^2$  so that the canonical realization of  $\alpha$  in  $\Sigma$  gives a realization in  $S^2$ . By the classification of compact surfaces, it suffices to prove that the intersection form  $B : H_1(\Sigma) \times H_1(\Sigma) \rightarrow \mathbb{Z}$  takes only even values. By Formula 3.5.1 and Lemma 4.2.1, we need only to prove that the expressions for the values of  $B$  appearing there are even. For all  $i \in E$ , we have  $n(e_i) \equiv \#(w_i) \equiv 0 \pmod{2}$  by the condition (i). Let  $i, j$  be distinct elements of  $E$ . If the arrows  $e_i, e_j \in \text{arr}(\alpha)$  are unlinked, then  $B([e_i], [e_j]) \equiv \#(w_i \cap w_j) \equiv 0 \pmod{2}$  by Formula 5.3.2 and the condition (ii). Suppose that  $e_i, e_j$  are linked. Assume first that the endpoints of  $e_i = (a_i, b_i), e_j = (a_j, b_j)$  lie on  $S$  in the cyclic order  $a_i, a_j, b_i, b_j$ . If  $i, j \in X$  (resp.  $i, j \in Y$ ), then both numbers  $a_i, a_j \in \{1, 2, \dots, 2m\}$  are odd (resp. even) so that  $q(e_i, e_j) \equiv a_i - a_j \equiv 0 \pmod{2}$ . By (iii), we have  $\#(w_i \cap w_j) \equiv 1 \pmod{2}$ . By Formula 5.3.3,  $B([e_i], [e_j])$  is even. If  $i, j$  lie in different subsets of the bipartition, then one of the numbers  $a_i, a_j$  is even and the other one is odd so that  $q(e_i, e_j) \equiv a_i - a_j \equiv 1 \pmod{2}$ . By (iii),  $\#(w_i \cap w_j) \equiv 0 \pmod{2}$  and by Formula 5.3.3,  $B([e_i], [e_j])$  is even. The case where the endpoints of  $e_i, e_j$  lie on  $S$  in the cyclic order  $a_i, b_j, b_i, a_j$  follows from the previous one using the skew-symmetry of  $B$ .  $\square$

**Corollary 5.3.2.** (*P. Rosenstiehl [Ro]*) *A Gauss word  $w$  in an alphabet  $E$  is realizable by a closed curve in  $S^2$  if and only if  $w$  satisfies the conditions (i), (ii) of Theorem 5.3.1 and  $E$  admits a bipartition compatible with  $w$ .*

We say that two closed curves  $\omega, \omega' : S^1 \rightarrow \Sigma$  on a surface  $\Sigma$  are *homeomorphic* if there is a homeomorphism  $\varphi : \Sigma \rightarrow \Sigma$  such that  $\omega' = \varphi \circ \omega$ . Here  $\varphi$  is not required to preserve orientation in  $\Sigma$ .

**Theorem 5.3.3.** *Two closed curves on  $S^2$  yield the same pair (a Gauss word, a bipartition of the alphabet) if and only if these curves are homeomorphic.*

*Proof.* We first prove that homeomorphic closed curves on  $S^2$  give rise to the same Gauss words and the same bipartitions. This is obvious if the homeomorphism  $S^2 \rightarrow S^2$  relating the curves is orientation-preserving. It remains to prove that the Gauss word and the bipartition associated with a closed curve  $\omega : S^1 \rightarrow S^2$  are preserved under reversal of orientation of  $S^2$ . By its very definition, the Gauss word does not depend on this orientation. It remains to check that the string  $\alpha$  underlying  $\omega$  and the inverse string  $\bar{\alpha}$  give rise to the same bipartition of the set  $\text{arr}(\alpha) = \text{arr}(\bar{\alpha})$ . Let  $e = (a, b), f = (c, d) \in \text{arr}(\alpha)$ . An easy count shows that the number of endpoints of  $\alpha$  in the interior of the arc  $bd \subset S^1$  is equal modulo 2 to the number of endpoints of  $\alpha$  in the interior of the arc  $ac \subset S^1$  plus  $n(e) + n(f)$ . Since  $\alpha$  underlies a curve on the 2-sphere,  $n(e) \equiv n(f) \equiv 0 \pmod{2}$ . This shows that  $e, f \in \text{arr}(\alpha)$  are equivalent in the sense of Section 5.2 if and only if the corresponding arrows  $(b, a), (d, c) \in \text{arr}(\bar{\alpha})$  are equivalent.

Conversely, suppose that two closed curves  $\omega, \omega' : S^1 \rightarrow S^2$  yield the same pair (a Gauss word in an alphabet  $E$ , a bipartition of  $E$ ). We shall show that  $\omega, \omega'$  are homeomorphic. Let  $\alpha, \alpha'$  be the underlying virtual strings of  $\omega, \omega'$ , respectively. Our assumption implies that we can label the arrows of  $\alpha, \alpha'$  with elements of  $E$  so that  $\alpha, \alpha'$  give rise to the same Gauss word and the same bipartition of  $E$ . Since  $\alpha, \alpha'$  determine the same Gauss word, we may assume that they coincide up to the choice of orientation of arrows. We claim that either  $\alpha = \alpha'$  or these strings are opposite on all arrows. Suppose that  $\alpha$  has an arrow  $(a, b)$  such that  $(b, a) \in \text{arr}(\alpha')$ . If  $(c, d)$  is another arrow of  $\alpha$ , then as in the previous paragraph, a simple count shows that the number of endpoints of  $\alpha$  lying inside the arc  $ac$  has the opposite parity to the number of endpoints of  $\alpha$  lying inside the arc  $bc$ . Since  $\alpha, \alpha'$  determine the same bipartition of the set of arrows, the pair  $(c, d)$  cannot be an arrow of  $\alpha'$ . Hence  $(d, c) \in \text{arr}(\alpha')$  which proves the claim above. (This claim is compatible with the fact that the set of orientations on  $m$  arrows has  $2^m$  elements while the set of bipartitions of  $E$  has  $2^{m-1}$  elements where  $m = \#(E)$ .) Thus  $\alpha'$  is homeomorphic either to  $\alpha$  or to its inverse  $\bar{\alpha}$ . The second case can be reduced to the first one by composing  $\omega$  with an orientation-reversing self-homeomorphism of  $S^2$ . Thus we can assume that  $\alpha, \alpha'$  are homeomorphic. As we know,  $\omega$  is a composition of the canonical realization  $\omega_\alpha : S^1 \rightarrow \Sigma_\alpha$  with an embedding  $\Sigma_\alpha \hookrightarrow S^2$  and similarly for  $\omega'$ . A homeomorphism  $\alpha \rightarrow \alpha'$  extends to a homeomorphism  $\varphi : \Sigma_\alpha \rightarrow \Sigma_{\alpha'}$  transforming  $\omega_\alpha$  into  $\omega_{\alpha'}$ . Note that all the components of

$S^2 - \Sigma_\alpha, S^2 - \Sigma_{\alpha'}$  are discs. Since any homeomorphism of circles extends to a homeomorphism of discs bounded by these circles,  $\varphi$  extends to a homeomorphism  $S^2 \rightarrow S^2$  transforming  $\omega$  into  $\omega'$ .  $\square$

Theorem 5.3.3 shows that the Gauss word and the bipartition associated with a closed curve on  $S^2$  is a full homeomorphism invariant. Theorems 5.3.1 and 5.3.3 give a complete combinatorial description of the set of homeomorphism classes of curves on  $S^2$  in terms of Gauss words and bipartitions.

**5.4. Examples.** The parity condition (i) of Theorem 5.3.1 was pointed out by Gauss who knew that from  $m = 5$  on it is not sufficient. He gave as examples the sequences 1231245345 and 1231435425. They satisfy (i) but not (ii). The word  $w = 123456214365$  satisfies (i) and (ii) but the alphabet  $\{1, 2, 3, 4, 5, 6\}$  does not admit a bipartition compatible with  $w$ . Indeed, the letters 1, 3, 5 of this alphabet are pairwise  $w$ -interlaced and  $\#(w_1 \cap w_3) = \#(w_1 \cap w_5) = \#(w_3 \cap w_5) = 2$ . The compatibility would imply that 1, 3, 5 belong to pairwise different subsets of the bipartition which is impossible.

**5.5. Irreducible Gauss words.** A Gauss word is *irreducible* if neither itself nor its circular permutations can be written as a concatenation of two non-empty Gauss words (in smaller alphabets). For example, the word 1212 in the alphabet  $\{1, 2\}$  is irreducible while the word 1221 is not. The next theorem shows that if an irreducible Gauss word allows a compatible bipartition of the alphabet then such a bipartition is unique.

**Theorem 5.5.1.** *For an irreducible Gauss word  $w$  in an alphabet  $E$ , there is at most one bipartition of  $E$  compatible with  $w$ .*

*Proof.* Consider the graph  $P$  whose vertices are elements of  $E$  and in which two vertices are connected by an edge if and only if they are  $w$ -interlaced. Since  $w$  is irreducible,  $P$  is connected. (If it were disconnected, then realizing  $w$  by a virtual string with core circle  $S^1 \subset \mathbb{C}$  and arrows represented by geometric vectors, one would easily observe that  $w$  is not irreducible.) If a bipartition  $E = X \cup Y$  is compatible with  $w$ , then knowing for an element of  $E$  whether it lies in  $X$  or  $Y$  we can determine this for its immediate neighbors in  $P$ . Proceeding along  $P$ , we eventually determine this for all elements of  $E$ . Hence there is at most one such bipartition of  $E$ .  $\square$

**Corollary 5.5.2.** *(C. H. Dowker and M. B. Thistlethwaite [DT]) Two closed curves in  $S^2$  realizing the same irreducible Gauss word are homeomorphic.*

This follows directly from Theorems 5.3.3 and 5.5.1.

**Corollary 5.5.3.** *If a Gauss word  $w$  in an alphabet  $E$  is obtained by concatenation of  $k \geq 1$  irreducible Gauss words, then the number of bipartitions of  $E$  compatible with  $w$  is either 0 or  $2^{k-1}$ .*

*Proof.* Denote the set of bipartitions of the alphabet compatible with  $w$  by  $B(w)$ . Let  $w$  be obtained by concatenation of two Gauss words  $w_1, w_2$  in disjoint alphabets  $E_1, E_2$ , respectively. Intersecting a bipartition of  $E = E_1 \cup E_2$  with  $E_1, E_2$  we obtain a mapping,  $j$ , from the set of bipartitions of  $E$  into the set of pairs (a bipartition of  $E_1$ , a bipartition of  $E_2$ ). This mapping is 2-to-1. Since the letters of  $E_1$  are not  $w$ -interlaced with the letters of  $E_2$ , we have  $j(B(w)) = B(w_1) \times B(w_2)$ . Hence  $\#(B(w)) = 2 \#(B(w_1)) \#(B(w_2))$ . This implies our claim by induction on  $k$ , the case  $k = 1$  being Theorem 5.5.1.  $\square$

**5.6. Remarks.** 1. If a closed curve  $\omega : S^1 \rightarrow \Sigma$  on a surface  $\Sigma$  is  $\mathbb{Z}/2\mathbb{Z}$ -homologically trivial, then the associated bipartition of the set of its double points  $\bowtie(\omega)$  admits a simple geometric interpretation. The homological triviality of  $\omega$  implies that the components of  $\Sigma - \omega(S^1)$  can be colored white or black so that the components of  $\Sigma - \omega(S^1)$  adjacent to the same arc in  $\omega(S^1) - \bowtie(\omega)$  from opposite sides have different colors. With each point  $x \in \bowtie(\omega)$  we associate the component of  $\Sigma - \omega(S^1)$  adjacent to  $x$  and lying between the two positive tangent vectors of  $\omega$  at  $x$ . It is easy to deduce from the definitions that two points of  $\bowtie(\omega)$  belong to the same subset of the bipartition determined by  $\omega$  if and only if the associated components of  $\Sigma - \omega(S^1)$  have the same color. In particular, for such  $\omega$ , the bipartition of  $\bowtie(\omega)$  determined by  $\omega$  does not depend on the orientation of  $\Sigma$ .

2. Theorem 5.3.3 can be generalized to certain curves on surfaces of arbitrary genus. Let us call a curve  $\omega : S^1 \rightarrow \Sigma$  on a surface  $\Sigma$  *special* if it is  $\mathbb{Z}/2\mathbb{Z}$ -homologically trivial and all the components of  $\Sigma - \omega(S^1)$  are discs. The proof of Theorem 5.3.3 shows that two special closed curves on a closed surface yield the same pair (a Gauss word, a bipartition of the alphabet) if and only if these curves are homeomorphic.

3. To study curves on  $\mathbb{R}^2$  one can involve an additional piece of combinatorial data which is a subset of the alphabet. Indeed, the complement of a curve in  $\mathbb{R}^2$  has one infinite region; the labels of the double points

adjacent to this region form a subset of the alphabet. It is easy to see that the Gauss word, the bipartition, and this subset form a full homeomorphism invariant of a closed curve on  $\mathbb{R}^2$ .

## 6. BASED SKEW-SYMMETRIC MATRICES

We introduce algebraic notions used in the sequel to define further homotopy invariants of strings.

**6.1. Definitions.** A *based skew-symmetric matrix over  $\mathbb{Z}$*  or shortly a *based matrix* is a triple  $(G, s, b)$  where  $G$  is a finite set,  $s \in G$ , and  $b : G^2 = G \times G \rightarrow \mathbb{Z}$  is a skew-symmetric mapping so that  $b(g, h) = -b(h, g)$  for all  $g, h \in G$ . In particular,  $b(g, g) = 0$  for all  $g \in G$ .

Two based matrices  $(G, s, b)$  and  $(G', s', b')$  are *isomorphic* if there is a bijection  $G \rightarrow G'$  sending  $s$  into  $s'$  and transforming  $b$  into  $b'$ . To specify the isomorphism class of a based matrix  $(G, s, b)$ , it suffices to specify the matrix  $(b(g, h))_{g, h \in G}$  where it is understood that the first column and row correspond to  $s$ . In this way every skew-symmetric square matrix over  $\mathbb{Z}$  determines a based matrix.

We call an element  $g \in G - \{s\}$  *annihilating* (with respect to  $b$ ) if  $b(g, h) = 0$  for all  $h \in G$ . We call  $g \in G - \{s\}$  a *core element* if  $b(g, h) = b(s, h)$  for all  $h \in G$ . We call two elements  $g_1, g_2 \in G - \{s\}$  *complementary* if  $b(g_1, h) + b(g_2, h) = b(s, h)$  for all  $h \in G$ . A based matrix  $(G, s, b)$  is *primitive* if it has no annihilating elements, no core elements, and no complementary pairs of elements. An example of a primitive based matrix is provided by the *trivial based matrix*  $(G, s, b)$  where  $G$  consists of only one element  $s$  and  $b(s, s) = 0$ .

**6.2. Equivalence of based matrices.** We define three operations  $M_1, M_2, M_3$  on based matrices, called *elementary extensions*. They add to a based matrix  $(G, s, b)$  an annihilating element, a core element, and a pair of complementary elements, respectively. More precisely,  $M_1$  transforms  $(G, s, b)$  into the (unique) based matrix  $(\bar{G} = G \amalg \{g\}, s, \bar{b})$  such that  $\bar{b} : \bar{G} \times \bar{G} \rightarrow \mathbb{Z}$  extends  $b$  and  $\bar{b}(g, h) = 0$  for all  $h \in \bar{G}$ . The move  $M_2$  transforms  $(G, s, b)$  into the (unique) based matrix  $(\hat{G} = G \amalg \{g\}, s, \hat{b})$  such that  $\hat{b} : \hat{G} \times \hat{G} \rightarrow \mathbb{Z}$  extends  $b$  and  $\hat{b}(g, h) = \hat{b}(s, h)$  for all  $h \in \hat{G}$ . The move  $M_3$  transforms  $(G, s, b)$  into a based matrix  $(\hat{G} = G \amalg \{g_1, g_2\}, s, \hat{b})$  where  $\hat{b} : \hat{G} \times \hat{G} \rightarrow \mathbb{Z}$  is any skew-symmetric map extending  $b$  and such that  $\hat{b}(g_1, h) + \hat{b}(g_2, h) = b(s, h)$  for all  $h \in \hat{G}$ . It is clear that a based matrix  $(G, s, b)$  is primitive if and only if it cannot be obtained from another based matrix by an elementary extension.

Two based matrices are *homologous* if one can be obtained from the other by a finite sequence of elementary extensions  $M_1, M_2, M_3$ , the inverse transformations, and isomorphisms. The homology is an equivalence relation on the set of based matrices. A simple homology invariant of a based matrix  $T = (G, s, b)$  is provided by the 1-variable polynomial

$$u_T(t) = \sum_{e \in G, b(e, s) \neq 0} \text{sign}(b(e, s)) t^{|b(e, s)|}.$$

**Lemma 6.2.1.** *Every based matrix is obtained from a primitive based matrix by elementary extensions. Two homologous primitive based matrices are isomorphic.*

*Proof.* The first claim is obvious: eliminating annihilating elements, core elements, and complementary pairs of elements by the moves  $M_i^{-1}$  with  $i = 1, 2, 3$  we can transform any based matrix  $T$  into a primitive based matrix  $T_0$ . Then  $T$  is obtained from  $T_0$  by elementary extensions.

To prove the second claim, we need the following assertion:

(\*) a move  $M_i$  followed by  $M_j^{-1}$  yields the same result as an isomorphism, or a move  $M_k^{\pm 1}$ , or a move  $M_l^{-1}$  followed by  $M_l$  with  $k, l \in \{1, 2, 3\}$ .

This assertion will imply the second claim of the lemma. Indeed, suppose that two primitive based matrices  $T, T'$  are related by a finite sequence of transformations  $M_1^{\pm 1}, M_2^{\pm 1}, M_3^{\pm 1}$  and isomorphisms. An isomorphism of based matrices followed by  $M_i^{\pm 1}$  can be also obtained as  $M_i^{\pm 1}$  followed by an isomorphism. Therefore all isomorphisms in our sequence can be accumulated at the end. The claim (\*) implies that  $T, T'$  can be related by a finite sequence of moves consisting of several moves of type  $M_i^{-1}$  followed by several moves of type  $M_i$  and isomorphisms. However, since  $T$  is primitive we cannot apply to it a move of type  $M_i^{-1}$ . Hence there are no such moves in our sequence. Similarly, since  $T'$  (and any isomorphic based matrix) is primitive, it cannot be obtained by an application of  $M_i$ . Therefore our sequence consists solely of isomorphisms so that  $T$  is isomorphic to  $T'$ .

Let us now prove (\*). We have to consider nine cases depending on the values  $i, j \in \{1, 2, 3\}$ .

For  $i, j \in \{1, 2\}$ , the move  $M_i$  on a based matrix  $(G, s, b)$  adds one element  $g$  and then  $M_j^{-1}$  removes one element  $g' \in G \amalg \{g\}$ . If  $g' = g$ , then  $M_j^{-1} \circ M_i$  is the identity. If  $g' \neq g$ , then  $g' \in G$  is annihilating (resp. core) for  $j = 1$  (resp.  $j = 2$ ). The transformation  $M_j^{-1} \circ M_i$  can be achieved by first applying  $M_j^{-1}$  that removes  $g'$  and then applying  $M_i$  that adds  $g$ .

Let  $i = 1, j = 3$ . The move  $M_i$  on  $(G, s, b)$  adds an annihilating element  $g$  and  $M_j^{-1}$  removes two complementary elements  $g_1, g_2 \in G \amalg \{g\}$ . If  $g_1 \neq g$  and  $g_2 \neq g$ , then  $g_1, g_2 \in G$  and  $M_j^{-1} \circ M_i$  can be achieved by first removing  $g_1, g_2$  and then adding  $g$ . If  $g_1 = g$ , then  $g_2$  is a core element of  $G$  and  $M_j^{-1} \circ M_i$  is the move  $M_2^{-1}$  removing  $g_2$ . The case  $g_2 = g$  is similar.

Let  $i = 2, j = 3$ . The move  $M_i$  on  $(G, s, b)$  adds a core element  $g$  and  $M_j^{-1}$  removes two complementary elements  $g_1, g_2 \in G \amalg \{g\}$ . If  $g_1 \neq g$  and  $g_2 \neq g$ , then  $M_j^{-1} \circ M_i$  can be achieved by first removing  $g_1, g_2$  and then adding  $g$ . If  $g_1 = g$ , then  $g_2 \in G$  is an annihilating element of  $G$  and  $M_j^{-1} \circ M_i$  is the move  $M_1^{-1}$  removing  $g_2$ . The case  $g_2 = g$  is similar.

Let  $i = 3, j = 1$ . The move  $M_i$  on  $(G, s, b)$  adds two complementary elements  $g_1, g_2$  and  $M_j^{-1}$  removes an annihilating element  $g \in G \amalg \{g_1, g_2\}$ . If  $g \neq g_1$  and  $g \neq g_2$ , then  $g \in G$  and  $M_j^{-1} \circ M_i$  can be achieved by first removing  $g$  and then adding  $g_1, g_2$ . If  $g = g_1$ , then  $g_2$  is a core element of  $G \amalg \{g_2\}$  and  $M_j^{-1} \circ M_i = M_2$ . The case  $g = g_2$  is similar.

Let  $i = 3, j = 2$ . The move  $M_i$  on  $(G, s, b)$  adds two complementary elements  $g_1, g_2$  and  $M_j^{-1}$  removes a core element  $g \in G \amalg \{g_1, g_2\}$ . If  $g \neq g_1$  and  $g \neq g_2$ , then  $g \in G$  and  $M_j^{-1} \circ M_i$  can be achieved by first removing  $g$  and then adding  $g_1, g_2$ . If  $g = g_1$ , then  $g_2$  is an annihilating element of  $G \amalg \{g_2\}$  and  $M_j^{-1} \circ M_i = M_1$ . The case  $g = g_2$  is similar.

Let  $i = j = 3$ . The move  $M_i$  on  $(G, s, b)$  adds two complementary elements  $g_1, g_2$  and  $M_j^{-1}$  removes two complementary elements  $g'_1, g'_2 \in G \amalg \{g_1, g_2\}$ . If these two pairs are disjoint, then  $M_j^{-1} \circ M_i$  can be achieved by first removing  $g'_1, g'_2 \in G$  and then adding  $g_1, g_2$ . If these two pairs coincide, then  $M_j^{-1} \circ M_i$  is the identity. It remains to consider the case where these pairs have one common element, say  $g'_1 = g_1$ , while  $g'_2 \neq g_2$ . Then  $g'_2 \in G$  and for all  $h \in G$ ,

$$\hat{b}(g_2, h) = b(s, h) - \hat{b}(g_1, h) = b(s, h) - \hat{b}(g'_1, h) = \hat{b}(g'_2, h) = b(g'_2, h).$$

Therefore the move  $M_j^{-1} \circ M_i$  gives a based matrix isomorphic to  $(G, s, b)$ . The isomorphism  $G \rightarrow (G - \{g'_2\}) \cup \{g_2\}$  is the identity on  $G - \{g'_2\}$  and sends  $g'_2$  into  $g_2$ .  $\square$

Lemma 6.2.1 implies that each based matrix  $T = (G, s, b)$  is homologous to a primitive based matrix  $T_0 = (G_0, s, b_0)$  unique up to isomorphism. This reduces the classification of homology classes of based matrices to an isomorphism classification of primitive based matrices. Note that we can choose  $T_0$  in its isomorphism class so that  $G_0 \subset G$  and  $b_0$  is the restriction of  $b$  to  $G_0 \times G_0$ .

Each isomorphism invariant  $v$  of primitive based matrices extends to a homology invariant of based matrices by  $v(T) = v(T_0)$ . The most important numerical invariant of a primitive based matrix  $(G, s, b)$  is the number  $\#(G)$ . It is easy to define further invariants of primitive based matrices. For instance, for  $k \in \mathbb{Z}$ , we can set

$$v_k(G, s, b) = \#\{g \in G \mid b(g, s) = k\}.$$

Similarly, for  $k \in \mathbb{Z}$  and a finite set of integers  $A$  endowed with non-negative multiplicities, set

$$v_{k,A}(G, s, b) = \#\{g \in G \mid b(g, s) = k \text{ and } \{b(g, h)\}_{h \in G - \{s\}} = A\},$$

where the latter equality is understood as an equality of sets with multiplicities. Clearly,  $v_k = \sum_A v_{k,A}$ .

**6.3. Remark.** The moves  $M_1, M_2, M_3$  on based matrices are not independent. It is easy to present  $M_2$  as a composition of  $M_3$  with  $M_1^{-1}$ .

**6.4. Genus of based matrices.** We define a numerical invariant of a based matrix  $T = (G, s, b)$  called its *genus* and denoted  $\sigma(T)$ . For subsets  $X, Y \subset G$ , set  $b(X, Y) = \sum_{g \in X, h \in Y} b(g, h) \in \mathbb{Z}$ . Clearly,  $b(X, Y) = -b(Y, X)$  and  $b(X, X) = b(\emptyset, X) = 0$  for all  $X, Y \subset G$ . A *regular partition* of  $G$  is a splitting of  $G$  as a union of disjoint (possibly empty) subsets  $\{X_i\}_i$  such that  $\#(X_i) \leq 2$  for all  $i$  and one of  $X_i$  is the one-element set  $\{s\}$ . The *matrix* of the regular partition  $x = \{X_i\}_i$  is the matrix  $(b(X_i, X_j))_{i,j}$ . This is a skew-symmetric square matrix over  $\mathbb{Z}$ . Its rank is an even number; let  $\sigma(x)$  denote half of this rank. By definition,  $\sigma(T) = \min_x \sigma(x)$  where  $x$  runs over all regular partitions of  $G$ . Extending  $b$  by linearity to the lattice generated by  $G$  and



identifying subsets of  $G$  with vectors in this lattice whose coordinates are 0 or 1, we can interpret  $\sigma(T)$  as half the minimal rank of the restriction of  $b$  to the sublattices arising from regular partitions of  $G$ .

Note that  $\sigma(T) \geq 0$  and  $\sigma(T) = 0$  if and only if  $G$  has a regular partition  $\{X_i\}_i$  such that  $b(X_i, X_j) = 0$  for all  $i, j$ . In the latter case we say that  $T$  is *hyperbolic*. It is easy to see that if  $T$  is hyperbolic, then  $u_T = 0$ .

The key property of the genus  $\sigma(T)$  is contained in the following lemma.

**Lemma 6.4.1.** *The genus of a based matrix is a homology invariant.*

*Proof.* By Remark 6.3, it suffices to prove that  $\sigma(T) = \sigma(T')$  for any based matrix  $T' = (G', s, b')$  obtained from a based matrix  $T = (G, s, b)$  by a move  $M_i$  with  $i = 1, 3$ . The set  $G' - G$  consists of one element if  $i = 1$  and of two elements if  $i = 3$ . Pick a regular partition  $x = \{X_i\}_i$  of  $G$  such that  $\sigma(T) = \sigma(x)$ . Consider the regular partition  $x' = (G' - G) \cup \{X_i\}_i$  of  $G'$ . Its matrix is obtained from the one of  $x$  by adjoining a row and a column. For  $i = 1$ , these row and column are zero so that  $\sigma(x) = \sigma(x')$ . For  $i = 3$ , we have  $b(G' - G, Y) = b(\{s\}, Y)$  for all  $Y \subset G$ . Since one of the sets  $X_i$  equals  $\{s\}$ , we again obtain  $\sigma(x) = \sigma(x')$ . Hence  $\sigma(T') \leq \sigma(x') = \sigma(x) = \sigma(T)$ .

To prove the opposite inequality, pick a regular partition  $x' = \{X_i\}_i$  of  $G'$  such that  $\sigma(T') = \sigma(x')$ . Consider first the case  $i = 1$ . One of the sets  $X_i$  contains the 1-element set  $G' - G$ . We replace this  $X_i$  by  $X_i - (G' - G)$  and keep all the other  $X_i$ . This gives a regular partition  $x$  of  $G$  whose matrix coincides with the matrix of  $x'$ . Hence  $\sigma(T) \leq \sigma(x) = \sigma(x') = \sigma(T')$ . Let now  $i = 3$ . If one of the sets  $X_i$  is equal to  $G' - G = \{g_1, g_2\}$ , then removing this  $X_i$  from  $x'$  we obtain a regular partition  $x$  of  $G$ . As in the previous paragraph,  $\sigma(x) = \sigma(x')$ . Hence  $\sigma(T) \leq \sigma(x) = \sigma(x') = \sigma(T')$ . Suppose that the elements  $g_1, g_2$  of  $G' - G$  belong to different subsets, say  $X_1, X_2$ , of the partition  $x'$ . Then the sets  $X_i$  with  $i \neq 1, 2$  and  $X = (X_1 \cup X_2) - \{g_1, g_2\}$  form a regular partition of  $G$ . Let  $X_0$  be the term of the partitions  $x$  and  $x'$  equal to  $\{s\}$ . For any  $Y \subset G$ , we have

$$b(X, Y) = b(X_1, Y) + b(X_2, Y) - b(g_1, Y) - b(g_2, Y) = b(X_1, Y) + b(X_2, Y) - b(X_0, Y).$$

Therefore the skew-symmetric bilinear form determined by the matrix of  $x$  is induced from the skew-symmetric bilinear form determined by the matrix of  $x'$  via the linear map of the corresponding lattices sending the vectors  $X$  and  $\{X_i\}_{i \neq 1, 2}$  respectively to  $X_1 + X_2 - X_0$  and  $\{X_i\}_{i \neq 1, 2}$ . Hence  $\sigma(x) \leq \sigma(x')$  and  $\sigma(T) \leq \sigma(x) \leq \sigma(x') = \sigma(T')$ .  $\square$

**Corollary 6.4.2.** *A based matrix homologous to a hyperbolic based matrix is itself hyperbolic.*

**6.5. Transformations**  $T \mapsto -T$ ,  $T \mapsto T^-$ . We define two more operations on based matrices. For a based matrix  $T = (G, s, b)$ , set  $-T = (G, s, -b)$  and  $T^- = (G, s, b^-)$  where  $b^-(s, h) = -b(s, h)$ ,  $b^-(h, s) = -b(h, s)$  for all  $h \in G$  and  $b^-(g, h) = b(g, h) + b(s, g) - b(s, h)$  for all  $g, h \in G - \{s\}$ . The transformations  $T \mapsto -T$ ,  $T \mapsto T^-$  are commuting involutions on the set of based matrices. It is easy to check that they are compatible with homology and preserve the class of primitive based matrices. It follows from the definitions that  $(-T)_0 = -T_0$  and  $(T^-)_0 = (T_0)^-$ .

**6.6. Remark.** We can define the direct sum  $T_1 \oplus T_2$  of based matrices  $T_1 = (G_1, s_1, b_1)$ ,  $T_2 = (G_2, s_2, b_2)$  to be the based matrix  $(G, s, b)$  where  $G = (G_1 \amalg G_2)/s_1 = s_2$ , the element  $s \in G$  is defined by  $s = s_1 = s_2$ , and  $b : G^2 \rightarrow \mathbb{Z}$  extends both  $b_1$  and  $b_2$  and satisfies  $b(g_1, g_2) = 0$  for any  $g_1 \in G_1 - \{s_1\}$ ,  $g_2 \in G_2 - \{s_2\}$ . As an exercise, the reader may check that the direct sum of primitive based matrices is primitive and the based matrix  $T \oplus (-T)$  is hyperbolic for any  $T$ .

## 7. BASED MATRICES OF STRINGS

**7.1. The based matrix of a string.** With each virtual string  $\alpha$  we associate a based matrix  $T(\alpha) = (G, s, b)$ . Set  $G = G(\alpha) = \{s\} \amalg \text{arr}(\alpha)$ . To define  $b = b(\alpha) : G \times G \rightarrow \mathbb{Z}$ , we identify  $G$  with the basis  $s \cup \{[e]\}_{e \in \text{arr}(\alpha)}$  of  $H_1(\Sigma_\alpha)$ , see Section 4.2. The map  $b$  is obtained by restricting the homological intersection pairing  $H_1(\Sigma_\alpha) \times H_1(\Sigma_\alpha) \rightarrow \mathbb{Z}$  to  $G$ . It is clear that  $b$  is skew-symmetric. We can compute  $b$  combinatorially using Formula 3.5.1 and Lemma 4.2.1. In particular,  $b(e, s) = n(e)$  for all  $e \in \text{arr}(\alpha)$ .

The map  $b$  can be computed from any closed curve  $\omega$  realizing  $\alpha$  on a surface  $\Sigma$ . Indeed, such a curve is obtained from the canonical realization of  $\alpha$  in  $\Sigma_\alpha$  via an orientation-preserving embedding  $\Sigma_\alpha \hookrightarrow \Sigma$ . It remains to observe that such an embedding preserves intersection numbers and transforms the basis  $s \cup \{[e]\}_{e \in \text{arr}(\alpha)}$  of  $H_1(\Sigma_\alpha)$  into the subset  $[\omega], \{[\omega_x]\}_{x \in \mathfrak{X}(\omega)}$  of  $H_1(\Sigma)$ , cf. Section 3.5.

**Lemma 7.1.1.** *If two virtual strings are homotopic, then their based matrices are homologous.*

*Proof.* By Lemma 4.4.1 it is enough to show that if two closed curves  $\omega, \omega'$  on a surface  $\Sigma$  are homotopic, then the based matrices of their underlying strings are homologous. By the discussion in Section 2.3, it suffices to consider the case where  $\omega'$  is obtained from  $\omega$  by one of the local moves listed there.

If  $\omega'$  is obtained from  $\omega$  by adding a small curl, then  $\bowtie(\omega') = \bowtie(\omega) \cup \{y\}$  where  $y$  is a new crossing. Clearly  $[\omega'_y] = 0 \in H_1(\Sigma)$  or  $[\omega'_y] = [\omega'] = [\omega] \in H_1(\Sigma)$  depending on whether the curl lies on the right or on the left of  $\omega$ . Also  $[\omega'_x] = [\omega_x]$  for all  $x \in \bowtie(\omega)$ . Hence  $T(\beta)$  is obtained from  $T(\alpha)$  by  $M_1$  or  $M_2$ .

Suppose that  $\omega'$  is obtained from  $\omega$  by the move pushing a branch of  $\omega$  across another branch and creating two new double points  $y, z$ . Clearly,  $[\omega'_x] = [\omega_x]$  for all  $x \in \bowtie(\omega) \subset \bowtie(\omega')$ . It is easy to see that  $[\omega'_y] + [\omega'_z] = [\omega'] = [\omega] \in H_1(\Sigma)$ . Therefore  $T(\beta)$  is obtained from  $T(\alpha)$  by  $M_3$ .

If  $\omega'$  is obtained from  $\omega$  by pushing a branch of  $\omega$  across a double point, then the subsets  $[\omega], \{[\omega_x]_{x \in \bowtie(\omega)}\}$  and  $[\omega'], \{[\omega'_x]_{x \in \bowtie(\omega')}\}$  of  $H_1(\Sigma)$  coincide so that  $T(\alpha)$  is isomorphic to  $T(\beta)$ .  $\square$

**7.2. Homotopy invariants of strings from based matrices.** Every virtual string  $\alpha$  gives rise to a primitive based matrix  $T_0(\alpha)$  by  $T_0(\alpha) = (T(\alpha))_0$ . This is the only primitive based matrix (up to isomorphism) homologous to  $T(\alpha)$ . By Lemma 7.1.1, the based matrix  $T_0(\alpha) = (G_0, s, b_0)$  is a homotopy invariant of  $\alpha$ . This based matrix determines the polynomial  $u(\alpha)$  introduced in Section 3: it follows from Formulas 3.2.2 and 3.5.1 that  $u(\alpha) = u_{T(\alpha)} = u_{T_0(\alpha)}$ . The number  $\rho(\alpha) = \#(G_0) - 1$  is a useful homotopy invariant of  $\alpha$  which may be non-zero even when  $u(\alpha) = 0$ , cf. the examples below. Note that if  $\alpha$  is homotopically trivial, then  $T_0(\alpha)$  is a trivial based matrix and  $\rho(\alpha) = 0$ .

It follows from the definitions that  $T(\alpha^-) = (T(\alpha))^-$  and therefore  $T_0(\alpha^-) = (T_0(\alpha))^-$ . Similarly,  $T(\bar{\alpha}) = -(T(\alpha))^-$  and  $T_0(\bar{\alpha}) = -(T_0(\alpha))^-$ .

The based matrix  $T_0(\alpha) = (G_0, s, b_0)$  can be used to estimate the homotopy rank and the homotopy genus of  $\alpha$ . Namely,  $hr(\alpha) \geq \rho(\alpha)$  since any string homotopic to  $\alpha$  must have at least  $\rho(\alpha)$  arrows. Similarly,  $hg(\alpha) \geq (1/2)\text{rank } b_0$  where  $\text{rank } b_0$  is the rank of the integral matrix  $(b_0(g, h))_{g, h \in G_0}$ . Indeed, if  $\alpha'$  is a string homotopic to  $\alpha$  and  $T(\alpha') = (G', s', b')$ , then  $g(\alpha') = (1/2)\text{rank } b' \geq (1/2)\text{rank } b_0$  since the matrix of  $b'$  contains the matrix of  $b_0$  as a submatrix.

Combining the inequalities  $hr(\alpha) \geq \rho(\alpha)$ ,  $hg(\alpha) \geq (1/2)\text{rank } b_0$  with the obvious inequalities  $\text{rank } \alpha \geq hr(\alpha)$  and  $g(\alpha) \geq hg(\alpha)$ , we obtain that if  $T(\alpha)$  is primitive, then  $hr(\alpha) = \text{rank } \alpha$  and  $hg(\alpha) = g(\alpha)$ .

The next theorem gives an estimate for the slice genus  $sg(\alpha)$  of  $\alpha$  via  $\sigma(T(\alpha)) = \sigma(T_0(\alpha))$ .

**Theorem 7.2.1.** *For any string  $\alpha$ , we have  $\sigma(T(\alpha)) \leq 2sg(\alpha)$ .*

*Proof.* Set  $k = sg(\alpha)$ . We can present  $\alpha$  by a loop on the boundary of a handlebody  $H$  which bounds a (singular) surface of genus  $k$  in  $H$ . This loop is homotopic in  $\partial H$  to a loop

$$\omega = \prod_{i=1}^n r_i^+ m'_i (r_i^-)^{-1} \prod_{j=1}^k p_j^+ q_j^+ (p_j^-)^{-1} (q_j^-)^{-1}$$

where  $r_i^+, r_i^-, m_i, m'_i$  are as in the proof of Theorem 3.5.1 and the paths  $p_j^-, q_j^-$  on  $\partial H$  are obtained from paths  $p_j^+, q_j^+$  on  $\partial H$ , respectively, by slight pushing to the right. Choosing the paths  $p_j^\pm, q_j^\pm, r_i^\pm$  carefully, we can assume that they begin and end in a small disc  $V \subset \partial H - \cup_{i=1}^n m_i$  and have no crossings in  $V$ . Then the crossings of  $\omega$  split into pairs of points  $(y, z)$  arising when

- (a) the paths  $p_j^+, p_j^-$  meet one of the paths  $q_k^\pm$ ;
- (b) the paths  $p_j^+, p_j^-$  meet one of the paths  $p_k^\pm, r_i^\pm, m'_i$ ;
- (c) the paths  $q_j^+, q_j^-$  meet one of the paths  $q_k^\pm, r_i^\pm, m'_i$ ;
- (d) the paths  $r_i^+, r_i^-$  meet one of the paths  $r_l^\pm$ ;
- (e) the paths  $r_i^+, r_i^-$  meet one of the paths  $m'_l$ .

Such pairs  $(y, z)$  give rise to homology classes  $[\omega_y], [\omega_z] \in H_1(\partial H)$  whose sum can be explicitly computed. Set  $s = [\omega] = [m_1] + \dots + [m_n] \in H_1(\partial H)$  and let  $[p_j], [q_j] \in H_1(\partial H, V) = H_1(\partial H)$  be the homology classes of the paths  $p_j^\pm, q_j^\pm$ , respectively. In the case (a),  $[\omega_y] + [\omega_z] = s \pm [q_j]$  if  $k \neq j$  and  $[\omega_y] + [\omega_z] = s \pm (s - [q_j])$  if  $k = j$ . In the case (b),  $[\omega_y] + [\omega_z] = s \pm [q_j]$ . In the case (c),  $[\omega_y] + [\omega_z] = s \pm [p_j]$ . In the case (d),  $[\omega_y] + [\omega_z] = s \pm [m_i]$ . In the case (e),  $[\omega_y] + [\omega_z] = s \pm [m_i]$  if  $l \neq i$  and  $[\omega_y] + [\omega_z] = s \pm (s - [m_i])$  if  $l = i$ . These computations show that the sublattice of  $H_1(\partial H)$  generated by such sums  $[\omega_y] + [\omega_z]$  is contained in the sublattice of  $H_1(\partial H)$  generated by  $2k + n$  elements  $[p_j], [q_j], [m_i]$ . Since  $B([m_i], [m_l]) = 0$  for all  $i, l$ , the restriction of the intersection form  $B : H_1(\partial H) \times H_1(\partial H) \rightarrow \mathbb{Z}$  to the latter (and hence to the former)

sublattice has rank  $\leq 4k$ . Therefore for the underlying string  $\alpha_\omega$  of  $\omega$  we have  $\sigma(T(\alpha_\omega)) \leq 2k$ . Since  $\alpha_\omega$  is homotopic to  $\alpha$ , their based matrices are homologous. By Lemma 6.4.1,  $\sigma(T(\alpha)) = \sigma(T(\alpha_\omega)) \leq 2k$ .  $\square$

**Corollary 7.2.2.** *For a slice string  $\alpha$ , the based matrices  $T(\alpha)$  and  $T_0(\alpha)$  are hyperbolic.*

**7.3. Applications.** (1) The based matrix  $T(\alpha_{p,q})$  of the string  $\alpha_{p,q}$  with  $p, q \geq 1$  was computed in Section 4.3. It is easy to check that except in the case  $p = q = 1$ , this based matrix is primitive. Thus  $T_0(\alpha_{p,q}) = T(\alpha_{p,q})$ ,  $hr(\alpha_{p,q}) = \text{rank } \alpha_{p,q} = p + q$  and  $hg(\alpha_{p,q}) = g(\alpha_{p,q})$  provided  $p \neq 1$  or  $q \neq 1$ . In particular,  $\alpha_{p,p}$  is a homotopically non-trivial string with zero  $u$ -polynomial for all  $p > 1$ .

(2) We show that the product of strings defined in Section 3.4 does not induce a well-defined operation on the set of homotopy classes of strings. To this end, we exhibit a homotopically trivial string whose product with itself is not homotopically trivial. Consider the permutation  $\sigma = (12)(34)$  on the set  $\{1, 2, 3, 4\}$  permuting 1 with 2 and 3 with 4. Consider the rank 4 string  $\alpha_\sigma$ , as defined in Section 3.3.2. Drawing a picture of  $\alpha_\sigma$ , one observes that it is a product of two copies of  $\alpha_{(12)}$ . The latter string is homotopically trivial since it is obtained from a trivial string by the homotopy move (b)<sub>s</sub>. The based matrix  $T(\alpha_\sigma)$  can be explicitly computed, cf. Section 4.3.2. It is determined by the following skew-symmetric matrix:

$$\begin{bmatrix} 0 & -1 & 1 & -1 & 1 \\ 1 & 0 & 1 & -1 & 1 \\ -1 & -1 & 0 & -1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ -1 & -1 & -1 & -1 & 0 \end{bmatrix}.$$

It is easy to check that this based matrix is primitive. Hence  $\alpha_\sigma$  is not homotopically trivial. Moreover, it is not homotopic to a string with  $< 4$  arrows.

(3) We prove that the involution  $\alpha \mapsto \overline{\alpha}$  acts non-trivially on the set of homotopy classes of strings. Consider the permutation  $\sigma = (134)(2)$  on the set  $\{1, 2, 3, 4\}$  sending 1 to 3, 3 to 4, 4 to 1, and 2 to 2. Drawing the string  $\alpha_\sigma$  we obtain that  $\overline{\alpha_\sigma} = \alpha_\tau$  where  $\tau$  is the permutation  $(124)(3)$ . The based matrices  $T(\alpha_\sigma)$  and  $T(\alpha_\tau)$  can be explicitly computed. They are determined by the following skew-symmetric matrices:

$$\begin{bmatrix} 0 & -2 & 0 & -1 & 3 \\ 2 & 0 & 1 & 0 & 3 \\ 0 & -1 & 0 & 0 & 2 \\ 1 & 0 & 0 & 0 & 1 \\ -3 & -3 & -2 & -1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & -1 & -2 & 0 & 3 \\ 1 & 0 & -1 & 1 & 3 \\ 2 & 1 & 0 & 1 & 2 \\ 0 & -1 & -1 & 0 & 1 \\ -3 & -3 & -2 & -1 & 0 \end{bmatrix}.$$

The based matrices  $T(\alpha_\sigma)$  and  $T(\alpha_\tau)$  are not isomorphic; this is clear for instance from the fact that the first matrix has a row with three zeros while the second matrix does not have such a row. It is clear also that these based matrices are primitive. By Lemma 6.2.1, they are not homologous. Hence  $\alpha_\sigma$  is not homotopic to  $\alpha_\tau = \overline{\alpha_\sigma}$ .

**7.4. Remark.** For open strings one can define a refined version of based matrices incorporating the splitting  $\text{arr} = \text{arr}^+ \cup \text{arr}^-$ . A *refined based matrix* is a based matrix  $(G, s, b)$  endowed with a splitting of  $G - \{s\}$  as a union of disjoint subsets  $G^+$  and  $G^-$ . The moves  $M_1, M_2, M_3$  on refined based matrices are defined as above with  $\overline{G}^+ = G^+ \cup \{g\}$ ,  $\overline{G}^- = G^-$ ,  $\tilde{G}^+ = G^+$ ,  $\tilde{G}^- = G^- \cup \{g\}$ , and  $\hat{G}^+ = G^+ \cup \{g_1\}$ ,  $\hat{G}^- = G^- \cup \{g_2\}$ . We leave further details to the reader.

## 8. LIE COBRACKET FOR STRINGS

In this section we introduce a Lie cobracket in the free module generated by homotopy classes of strings. This induces a Lie bracket in the module of homotopy invariants of strings.

Throughout the section, we fix a commutative ring  $R$  with unit.

**8.1. Lie coalgebras.** We recall here the notion of a Lie coalgebra dual to the one of a Lie algebra. To this end, we first reformulate the notion of a Lie algebra. For an  $R$ -module  $L$ , denote by  $\text{Perm}_L$  the permutation  $x \otimes y \mapsto y \otimes x$  in  $L^{\otimes 2} = L \otimes L$  and by  $\tau_L$  the permutation  $x \otimes y \otimes z \mapsto z \otimes x \otimes y$  in  $L^{\otimes 3} = L \otimes L \otimes L$ . Here and below  $\otimes = \otimes_R$ . A Lie algebra over  $R$  is an  $R$ -module  $L$  endowed with an  $R$ -homomorphism (the Lie bracket)  $\theta : L^{\otimes 2} \rightarrow L$  such that  $\theta \circ \text{Perm}_L = -\theta$  (antisymmetry) and

$$\theta \circ (\text{id}_L \otimes \theta) \circ (\text{id}_{L^{\otimes 3}} + \tau_L + \tau_L^2) = 0 \in \text{Hom}_R(L^{\otimes 3}, L)$$

(the Jacobi identity). Dually, a Lie coalgebra over  $R$  is an  $R$ -module  $A$  endowed with an  $R$ -homomorphism (the Lie cobracket)  $\nu : A \rightarrow A^{\otimes 2}$  such that  $\text{Perm}_A \circ \nu = -\nu$  and

$$(8.1.1) \quad (\text{id}_{A^{\otimes 3}} + \tau_A + \tau_A^2) \circ (\text{id}_A \otimes \nu) \circ \nu = 0 \in \text{Hom}_R(A, A^{\otimes 3}).$$

A Lie coalgebra  $(A, \nu)$  gives rise to the *dual Lie algebra*  $A^* = \text{Hom}_R(A, R)$  where the Lie bracket  $A^* \otimes A^* \rightarrow A^*$  is the homomorphism dual to  $\nu$ . For  $u, v \in A^*$ , the value of  $[u, v] \in A^*$  on  $x \in A$  is computed by

$$[u, v](x) = \sum_i u(x_i^{(1)}) v(x_i^{(2)}) \in R$$

for any (finite) expansion  $\nu(x) = \sum_i x_i^{(1)} \otimes x_i^{(2)} \in A \otimes A$ .

A *homomorphism* of Lie coalgebras  $(A, \nu) \rightarrow (A', \nu')$  is an  $R$ -linear homomorphism  $\psi : A \rightarrow A'$  such that  $(\psi \otimes \psi)\nu(a) = \nu'\psi(a)$  for all  $a \in A$ . It is clear that the dual homomorphism  $\psi^* : (A')^* \rightarrow A^*$  is a homomorphism of Lie algebras.

**8.2. Lie coalgebra of strings.** Let  $\mathcal{S}$  be the set of homotopy classes of virtual strings and let  $\mathcal{S}_0 \subset \mathcal{S}$  be its subset formed by the homotopically non-trivial classes. Let  $\mathcal{A}_0 = \mathcal{A}_0(R)$  be the free  $R$ -module freely generated by  $\mathcal{S}_0$ . We shall provide  $\mathcal{A}_0$  with the structure of a Lie coalgebra.

We begin with notation. For a string  $\alpha$ , let  $\langle \alpha \rangle$  denote its class in  $\mathcal{S}_0$  if  $\alpha$  is homotopically non-trivial and set  $\langle \alpha \rangle = 0 \in \mathcal{A}_0$  if  $\alpha$  is homotopically trivial. For an arrow  $e = (a, b)$  of a string  $\alpha$ , denote by  $\alpha_e^1$  the string obtained from  $\alpha$  by removing all arrows except those with both endpoints in the interior of the arc  $ab$ . (In particular,  $e$  is removed.) Similarly, denote by  $\alpha_e^2$  the string obtained from  $\alpha$  by removing all arrows except those with both endpoints in the interior of  $ba$ . Set

$$(8.2.1) \quad \nu(\langle \alpha \rangle) = \sum_{e \in \text{arr}(\alpha)} \langle \alpha_e^1 \rangle \otimes \langle \alpha_e^2 \rangle - \langle \alpha_e^2 \rangle \otimes \langle \alpha_e^1 \rangle \in \mathcal{A}_0 \otimes \mathcal{A}_0.$$

**Lemma 8.2.1.** *The  $R$ -linear homomorphism  $\mathcal{A}_0 \rightarrow \mathcal{A}_0 \otimes \mathcal{A}_0$  given on the generators of  $\mathcal{A}_0$  by Formula 8.2.1 is a well-defined Lie cobracket.*

*Proof.* To show that  $\nu$  is well-defined we must verify that  $\nu(\langle \alpha \rangle)$  does not change under the homotopy moves  $(a)_s, (b)_s, (c)_s$  on  $\alpha$ . The arrow added by  $(a)_s$  contributes 0 to the cobracket by the definition of  $\langle \dots \rangle$ . The contribution of all the other arrows is preserved. Similarly, the two arrows added by  $(b)_s$  contribute opposite terms to the cobracket which is therefore preserved. Under  $(c)_s$ , all arrows contribute the same before and after the move.

The equality  $\text{Perm}_{\mathcal{A}_0} \circ \nu = -\nu$  is obvious. We now verify Formula 8.1.1. Let  $\alpha$  be a string with core circle  $S$ . We can expand  $(\text{id} \otimes \nu)(\nu(\langle \alpha \rangle))$  as a sum of expressions  $z(e, f)$  associated with ordered pairs of unlinked arrows  $e, f \in \text{arr}(\alpha)$ . Note that the endpoints of  $e, f$  split  $S$  into four arcs meeting only at their endpoints. The endpoints of  $e$  (resp.  $f$ ) bound one of these arcs, say  $x$  (resp.  $y$ ). The other two arcs form  $S - (x \cup y)$  and lie “between”  $e$  and  $f$ . Denote by  $\beta$  (resp.  $\gamma, \delta$ ) the string obtained from  $\alpha$  by removing all arrows except those with both endpoints in the interior of  $x$  (resp. of  $y$ , of  $S - (x \cup y)$ ). Set  $\varepsilon = +1$  if  $e$  and  $f$  are co-oriented, i.e., if their tails bound a component of  $S - (x \cup y)$ . It is easy to see that

$$z(e, f) = \varepsilon(\langle \beta \rangle \otimes \langle \delta \rangle \otimes \langle \gamma \rangle - \langle \beta \rangle \otimes \langle \gamma \rangle \otimes \langle \delta \rangle).$$

A direct computation using this formula gives

$$(\text{id}_{\mathcal{A}_0^{\otimes 3}} + \tau_{\mathcal{A}_0} + \tau_{\mathcal{A}_0}^2)(z(e, f) + z(f, e)) = 0.$$

Thus  $\text{id}_{\mathcal{A}_0^{\otimes 3}} + \tau_{\mathcal{A}_0} + \tau_{\mathcal{A}_0}^2$  annihilates  $(\text{id} \otimes \nu)(\nu(\langle \alpha \rangle))$ . Hence  $\nu$  is a Lie cobracket.  $\square$

**8.3. Lie coalgebra  $\mathcal{A}$  and Lie algebra  $\mathcal{A}^*$ .** Let  $\mathcal{A} = \mathcal{A}(R)$  be the free  $R$ -module freely generated by  $\mathcal{S}$ . Since  $\mathcal{S} = \mathcal{S}_0 \cup \{\Phi\}$  where  $\Phi \in \mathcal{S}$  is the homotopy class of a trivial string,  $\mathcal{A} = \mathcal{A}_0 \oplus R\Phi$ . The Lie cobracket  $\nu$  in  $\mathcal{A}_0$  extends to  $\mathcal{A}$  by  $\nu(\Phi) = 0$ .

The Lie cobrackets in  $\mathcal{A}_0$  and  $\mathcal{A}$  induce Lie brackets in  $\mathcal{A}_0^* = \text{Hom}_R(\mathcal{A}_0, R)$  and  $\mathcal{A}^* = \text{Hom}_R(\mathcal{A}, R)$ , respectively. Examples below show that these Lie cobrackets and Lie brackets are non-zero. Clearly,  $\mathcal{A}^* = \mathcal{A}_0^* \oplus R$  where the Lie bracket in  $R$  is zero.

The elements of  $\mathcal{A}^*$  bijectively correspond to maps  $\mathcal{S} \rightarrow R$ , i.e., to  $R$ -valued homotopy invariants of strings. Thus, such invariants form a Lie algebra.

**8.4. Examples.** (1) If  $\text{rank } \alpha \leq 6$ , then  $\nu(\langle \alpha \rangle) = 0$ . This follows from the fact that any string of rank  $\leq 2$  is homotopically trivial.

(2) For any  $p, q \geq 1$ , we have  $\nu(\langle \alpha_{p,q} \rangle) = 0$ .

(3) Consider the string  $\alpha = \alpha_\sigma$  of rank 7 where  $\sigma$  is the permutation (123)(4)(576) of the set  $\{1, 2, \dots, 7\}$ . It follows from the definitions that  $\nu(\langle \alpha \rangle) = \langle \alpha_{1,2} \rangle \otimes \langle \alpha_{2,1} \rangle - \langle \alpha_{2,1} \rangle \otimes \langle \alpha_{1,2} \rangle$ . As we know,  $\alpha_{1,2}$  and  $\alpha_{2,1}$  are homotopically non-trivial strings representing distinct generators of  $\mathcal{A}$ . Hence  $\nu(\langle \alpha \rangle) \neq 0$ .

(4) In generalization of the previous example pick any integers  $p, q, p', q' \geq 1$  such that  $p+q \geq 3, p'+q' \geq 3$ . Consider the string  $\alpha = \alpha_\sigma$  of rank  $m = p+q+p'+q'+1$  where  $\sigma$  is the permutation of the set  $\{1, 2, \dots, m\}$  defined by

$$\sigma(i) = \begin{cases} i+q, & \text{if } 1 \leq i \leq p \\ i-p, & \text{if } p < i \leq p+q \\ i, & \text{if } i = p+q+1 \\ i+q', & \text{if } p+q+1 < i \leq p+q+1+p' \\ i-p', & \text{if } p+q+1+p' < i \leq m. \end{cases}$$

It follows from the definitions that

$$\nu(\langle \alpha \rangle) = \langle \alpha_{p',q'} \rangle \otimes \langle \alpha_{p,q} \rangle - \langle \alpha_{p,q} \rangle \otimes \langle \alpha_{p',q'} \rangle.$$

Clearly,  $\nu(\langle \alpha \rangle) \neq 0$  unless  $p = p'$  and  $q = q'$ .

(5) Consider the numerical invariants  $u_1, u_2, \dots \in \mathcal{A}^*$  constructed in Section 3.1. For  $p, p' \geq 1$ , we compute the value of  $[u_p, u_{p'}] \in \mathcal{A}^*$  on the string  $\alpha = \alpha(p, p', q, q')$  defined in the previous example. Assume for concreteness that the numbers  $p, p', q, q'$  are pairwise distinct. Then

$$[u_p, u_{p'}](\alpha) = u_p(\alpha_{p',q'}) u_{p'}(\alpha_{p,q}) - u_p(\alpha_{p,q}) u_{p'}(\alpha_{p',q'}) = 0 - (-q)(-q') = -qq'.$$

Hence  $[u_p, u_{p'}] \neq 0$  for  $p \neq p'$ .

**8.5. Filtration of  $\mathcal{A}_0$ .** Assigning to a string its homotopy rang and homotopy genus (see Section 2.5) we obtain two maps  $hr, hg : \mathcal{S}_0 \rightarrow \mathbb{Z}$ . For  $r, g \geq 0$ , set

$$\mathcal{S}_{r,g} = \{\alpha \in \mathcal{S}_0 \mid hg(\alpha) \leq r, \quad hg(\alpha) \leq g\}.$$

This set is finite since there is only a finite number of strings of rank  $\leq r$ . The set  $\mathcal{S}_{r,g}$  generates a submodule of  $\mathcal{A}_0$  denoted  $\mathcal{A}_{r,g}$ . This submodule is a free  $R$ -module of rank  $\#(\mathcal{S}_{r,g})$ . Clearly,

$$(8.5.1) \quad \nu(\mathcal{A}_{r,g}) \subset \bigoplus_{p,q \geq 0, p+q < r} \mathcal{A}_{p,g} \otimes \mathcal{A}_{q,g} \subset \mathcal{A}_{r,g} \otimes \mathcal{A}_{r,g}.$$

Thus, each  $\mathcal{A}_{r,g}$  a Lie coalgebra. The inclusions  $\mathcal{A}_{r,g} \hookrightarrow \mathcal{A}_{r',g'}$  for  $r \leq r', g \leq g'$  make the family  $\{\mathcal{A}_{r,g}\}_{r,g}$  into a direct spectrum of Lie coalgebras. The equality  $\mathcal{A}_0 = \bigcup_{r,g} \mathcal{A}_{r,g}$  shows that  $\mathcal{A}_0 = \text{inj lim } \{\mathcal{A}_{r,g}\}$ .

The Lie cobracket in  $\mathcal{A}_{r,g}$  induces a Lie bracket in  $\mathcal{A}_{r,g}^* = \text{Hom}_R(\mathcal{A}_{r,g}, R)$ . Formula 8.5.1 implies that this Lie algebra is nilpotent. Restricting maps  $\mathcal{S}_0 \rightarrow R$  to  $\mathcal{S}_{r,g}$  we obtain a Lie algebra homomorphism  $\mathcal{A}_0^* \rightarrow \mathcal{A}_{r,g}^*$ . It is clear that  $\mathcal{A}_0^* = \text{proj lim } \{\mathcal{A}_{r,g}^*\}$ .

**8.6. Relations with Lie coalgebras of curves.** Let  $\Sigma$  be a connected surface and  $\hat{\pi}$  be the set of homotopy classes of closed curves on  $\Sigma$ . (It can be identified with the set of conjugacy classes in  $\pi = \pi_1(\Sigma)$ .) There is a map  $\psi : \hat{\pi} \rightarrow \mathcal{S}$  sending each homotopy class of curves into the homotopy class of the underlying strings. Clearly,  $\psi(\hat{\pi}) = \bigcup_{r,g} \mathcal{S}_{r,g}$  where  $g = g(\Sigma)$  is the genus of  $\Sigma$ . Observe that the mapping class group of  $\Sigma$  acts on  $\hat{\pi}$  in the obvious way and  $\psi$  factors through the projection of  $\hat{\pi}$  to the set of orbits of this action.

Let  $Z = Z(R)$  be the free  $R$ -module with basis  $\hat{\pi}$ . The map  $\psi : \hat{\pi} \rightarrow \mathcal{S}$  induces an  $R$ -linear homomorphism  $Z \rightarrow \mathcal{A}$  whose image is equal to  $\bigcup_r \mathcal{A}_{r,g}$ . Composing this homomorphism with the projection  $\mathcal{A} = \mathcal{A}_0 \oplus R\Phi \rightarrow \mathcal{A}_0$  we obtain an  $R$ -linear homomorphism  $\psi_0 : Z \rightarrow \mathcal{A}_0$ .

The author defined in [Tu2], Section 8 a structure of a Lie coalgebra in  $Z$ . (In fact  $Z$  is a Lie bialgebra, but we shall not use the Lie bracket in  $Z$ .) A direct comparison of the definitions yields the following.

**Lemma 8.6.1.** *The map  $\psi_0 : Z \rightarrow \mathcal{A}_0$  is a homomorphism of Lie coalgebras.*

Composing  $\psi_0$  with the inclusion  $\mathcal{A}_0 \hookrightarrow \mathcal{A}$  and dualizing we obtain a Lie algebra homomorphism  $\mathcal{A}^* \rightarrow Z^*$ .

**8.7. Applications.** We claim that the product of strings is not commutative even up to homotopy: there are strings  $\gamma, \delta$  such that a product of  $\gamma, \delta$  is not homotopic to a product of  $\delta, \gamma$ . Consider the string  $\alpha$  constructed in Example 8.4.3. Drawing a picture of  $\alpha$ , one observes that  $\alpha$  is a product of  $\delta = \alpha_{2,1}$  with a string,  $\gamma$ , of rank 4 obtained from  $\alpha_{1,2}$  by adding a “small” arrow. Since  $\gamma$  has a small arrow, it is easy to form a product of  $\gamma$  with  $\delta$  also having a small arrow. The resulting string,  $\beta$ , is homotopic to a string of rank 6. Hence  $\nu(\langle\beta\rangle) = 0$  whereas  $\nu(\langle\alpha\rangle) \neq 0$ . Therefore  $\alpha$  is not homotopic to  $\beta$ .

## 9. VIRTUAL STRINGS VERSUS VIRTUAL KNOTS

Virtual knots were introduced by L. Kauffman [Ka] as a generalization of classical knots. We relate them to virtual strings by showing that each virtual knot gives rise to a polynomial on virtual strings with coefficients in the ring  $\mathbb{Q}[z]$ . As a technical tool, we introduce a skein algebra of virtual knots and compute it in terms of strings.

**9.1. Virtual knots.** We define virtual knots in terms of arrow diagrams following [GPV]. An *arrow diagram* is a virtual string whose arrows are endowed with signs  $\pm$ . By the core circle and the endpoints of an arrow diagram, we mean the core circle and the endpoints of the underlying virtual string. The sign of an arrow  $e$  of an arrow diagram is denoted  $\text{sign}(e)$ . Homeomorphisms of arrow diagrams are defined as the homeomorphisms of the underlying strings preserving the signs of all arrows. The homeomorphism classes of arrow diagrams will be also called arrow diagrams.

We describe three moves  $(a)_{ad}$ ,  $(b)_{ad}$ ,  $(c)_{ad}$  on arrow diagrams where  $ad$  stands for “arrow diagram”. Let  $\alpha$  be an arrow diagram with core circle  $S$ . Pick two distinct points  $a, b \in S$  such that the (positively oriented) arc  $ab \subset S$  is disjoint from the set of endpoints of  $\alpha$ . The move  $(a)_{ad}$  adds to  $\alpha$  the arrow  $(a, b)$  with sign  $+$  or  $-$ . This move has two forms determined by the sign  $\pm$ . The move  $(b)_{ad}$  acts on  $\alpha$  as follows. Pick two arcs on  $S$  disjoint from each other and from the endpoints of  $\alpha$ . Let  $a, a'$  be the endpoints of the first arc (in an arbitrary order) and  $b, b'$  be the endpoints of the second arc. The move adds to  $\alpha$  two arrows  $(a, b)$  and  $(b', a')$  with opposite signs. This move has eight forms depending on the choice of the sign of  $(a, b)$ , two possible choices for  $a$ , and two possible choices for  $b$ . (This list of eight forms of  $(b)_{ad}$  contains two equivalent pairs so that in fact the move  $(b)_{ad}$  has only six forms.) The move  $(c)_{ad}$  applies to  $\alpha$  when  $\alpha$  has three arrows with signs  $((a^+, b), +), ((b^+, c), +), ((c^+, a), -)$  where  $a, a^+, b, b^+, c, c^+ \in S$  such that the arcs  $aa^+, bb^+, cc^+$  are disjoint from each other and from the other endpoints of  $\alpha$ . The move  $(c)_{ad}$  replaces these three arrows with the arrows  $((a, b^+), +), ((b, c^+), +), ((c, a^+), -)$ .

By definition, a *virtual knot* is an equivalence class of arrow diagrams with respect to the equivalence relation generated by the moves  $(a)_{ad}$ ,  $(b)_{ad}$ ,  $(c)_{ad}$  and homeomorphisms. Note that our set of moves is somewhat different from the one in [GPV] but generates the same equivalence relation (cf. below).

In the sequel the virtual knot represented by an arrow diagram  $D$  will be denoted  $[D]$ . A *trivial arrow diagram* having no arrows represents the *trivial virtual knot*.

Forgetting the signs of arrows, we can associate with any arrow diagram  $D$  its underlying virtual string  $\underline{D}$ . This induces a “forgetting” map  $K \mapsto \underline{K}$  from the set of virtual knots into the set of virtual strings. This map is surjective but not injective. The theory of virtual knots is considerably richer than the theory of virtual strings. For instance, the fundamental group of a virtual knot [Ka] allows to distinguish virtual knots with the same underlying strings.

**9.2. From knots to virtual knots.** Arrow diagrams are closely related to the standard knot diagrams on surfaces. An (oriented) knot diagram on an (oriented) surface  $\Sigma$  is a (generic oriented) closed curve on  $\Sigma$  such that at each its double point one of the branches of the curve passing through this point is distinguished. The distinguished branch is called an *overcrossing* while the second branch passing through the same point is called an *undercrossing*. A knot diagram on  $\Sigma = \Sigma \times \{0\}$  determines an (oriented) knot in  $\Sigma \times \mathbb{R}$  by pushing the overcrossings into  $\Sigma \times (0, \infty)$ .

Any knot diagram  $d$  gives rise to an arrow diagram  $D(d)$  as follows. First of all, the closed curve underlying  $d$  gives rise to a virtual string, see Section 2.2. We provide each arrow of this string with the sign of the corresponding double point of  $d$ . This sign is  $+$  (resp.  $-$ ) if the pair (a positive tangent vector to the overcrossing branch, a positive tangent vector to the undercrossing branch) is positive (resp. negative) with respect to the orientation of  $\Sigma$ . Our definition of the arrow diagram associated with  $d$  differs from the one in [GPV]: their arrow diagram is obtained from ours by reversing all arrows with sign  $-$ .

There is a canonical mapping from the set of isotopy classes of (oriented) knots in  $\Sigma \times \mathbb{R}$  into the set of virtual knots. It assigns to a knot  $K \subset \Sigma \times \mathbb{R}$  the virtual knot  $[D(d)]$  where  $d$  is a knot diagram on  $\Sigma$  presenting a knot in  $\Sigma \times \mathbb{R}$  isotopic to  $K$ . The virtual knot  $[D(d)]$  does not depend on the choice of  $d$ . This follows from the fact that two knot diagrams on  $\Sigma$  presenting isotopic knots in  $\Sigma \times \mathbb{R}$  can be obtained from each other by ambient isotopy in  $\Sigma$  and the Reidemeister moves. Recall the standard list of the Reidemeister moves: (1) a move adding a twist on the right (resp. left) of a branch; (2) a move pushing a branch over another branch and creating two crossings; (3) a move pushing a branch over a crossing. This list is redundant. In particular, the left move of type (1) can be presented as a composition of type (2) moves and the inverse to a right move of type (1). One move of type (3) together with moves of type (2) is sufficient to generate all moves of type (3) corresponding to various orientations on the branches (see, for instance, [Tu1], pp. 543–544). As the generating move of type (3) we take the move  $(c)^-$  described in Section 2.3. It remains to observe that the moves  $(a)_{ad}$ ,  $(b)_{ad}$ ,  $(c)_{ad}$  on arrow diagrams are exactly the moves induced by the right Reidemeister moves of type (1), the Reidemeister moves of type (2), and the move  $(c)^-$ .

**9.3. Skein algebra of virtual knots.** Let  $R = \mathbb{Q}[z]$  be the ring of polynomials in one variable  $z$  with rational coefficients. Consider the polynomial algebra  $R[\mathcal{K}]$  generated by the set of virtual knots  $\mathcal{K}$ . This is a commutative associative algebra with unit whose elements are polynomials in elements of  $\mathcal{K}$  with coefficients in  $R$ . We now introduce certain elements of  $R[\mathcal{K}]$  called “skein relations”.

Pick an arrow diagram  $D$  with core circle  $S$  and pick an arrow  $e = (a, b)$  of  $D$  with sign  $+$  (here  $a, b \in S$ ). Let  $D_e^-$  be the same arrow diagram with the sign of  $e$  changed to  $-$ . Let  $D'_e$  be the arrow diagram obtained from  $D$  by removing all arrows with at least one endpoint on the arc  $ba \subset S$ . Let  $D''_e$  be the arrow diagram obtained from  $D$  by removing all arrows with at least one endpoint on the arc  $ab \subset S$ . The skein relation corresponding to  $(D, e)$  is  $[D] - [D_e^-] - z[D'_e][D''_e] \in R[\mathcal{K}]$ .

The ideal of the algebra  $R[\mathcal{K}]$  generated by the trivial virtual knot and the skein relations (determined by all the pairs  $(D, e)$  as above) is called the *skein ideal*. The quotient of  $R[\mathcal{K}]$  by this ideal is called the *skein algebra of virtual knots* and denoted  $\mathcal{E}$ . The next theorem computes  $\mathcal{E}$  in terms of strings. Recall the set  $\mathcal{S}_0$  of non-trivial homotopy classes of virtual strings, cf. Section 8.2.

**Theorem 9.3.1.** *There is a canonical  $R$ -algebra isomorphism  $\nabla : \mathcal{E} \rightarrow R[\mathcal{S}_0]$  where  $R[\mathcal{S}_0]$  is the polynomial algebra generated by  $\mathcal{S}_0$ .*

This theorem allows us to associate with any virtual knot  $K$  a polynomial  $\nabla(K) \in R[\mathcal{S}_0]$ . It will be clear from the definitions that

$$\nabla(K) = \langle \underline{K} \rangle + \sum_{n \geq 2} z^{n-1} \nabla_n(K)$$

where  $\nabla_n(K)$  is a homogeneous element of  $\mathbb{Q}[\mathcal{S}_0]$  of degree  $n$  which is non-zero only for a finite set of  $n$ . Combining  $\nabla$  with homotopy invariants of strings we obtain invariants of virtual knots. For example, composing  $\nabla$  with the algebra homomorphism  $R[\mathcal{S}_0] \rightarrow R[t]$  sending the homotopy class of a string  $\alpha$  into the polynomial  $u(\alpha)(t)$ , we obtain an algebra homomorphism  $\mathcal{E} \rightarrow R[t] = \mathbb{Q}[z, t]$ . This gives a 2-variable polynomial invariant of virtual knots.

The constructions above can be applied to virtual knots derived from geometric knots in  $(\text{a surface}) \times \mathbb{R}$  as in Section 9.2. The resulting invariants are interesting only in the case when the genus of the surface is at least 2. This is due to the fact that the strings realized by curves on a surface of genus 0 or 1 are homotopically trivial.

Theorem 9.3.1 will be proven in the next section. Here we give an explicit expression for the value of  $\nabla$  on the generator  $[D] \in \mathcal{E}$  represented by an arrow diagram  $D$ . We need a few definitions. The endpoints of the arrows of  $D$  split the core circle of  $D$  into (oriented) arcs called the *edges* of  $D$ . Denote the set of edges of  $D$  by  $\text{edg}(D)$ . Each endpoint  $a$  of an arrow of  $D$  is adjacent to two edges  $a_-, a_+ \in \text{edg}(D)$ , respectively incoming and outgoing with respect to  $a$ . For an integer  $n \geq 1$ , an  $n$ -*labeling* of  $D$  is a map  $f : \text{edg}(D) \rightarrow \{1, 2, \dots, n\}$  satisfying the following condition: for any arrow  $e = (a, b)$  of  $D$ , either

- (i)  $f(a_+) = f(a_-)$ ,  $f(b_+) = f(b_-)$  or
- (ii)  $f(a_+) = f(b_-) \neq f(a_-) = f(b_+)$  and  $\text{sign}(f(a_-) - f(a_+)) = \text{sign}(e)$ .

The arrows  $e$  as in (ii) are called  *$f$ -cutting arrows*. The number of  $f$ -cutting arrows of  $D$  is denoted  $|f|$  and the number of  $f$ -cutting arrows of  $D$  with sign  $-1$  is denoted  $|f|_-$ . Note that the value of  $f$  on two adjacent edges  $a_-, a_+ \in \text{edg}(D)$  may differ only when  $a$  is an endpoint of an  $f$ -cutting arrow. Therefore

$|f| \geq \#f(\text{edg}(D)) - 1$ . For  $i = 1, \dots, n$ , let  $\underline{D}_{f,i}$  be the string obtained from  $D$  by removing all arrows except the arrows  $(a, b)$  with  $f(a_+) = f(a_-) = f(b_+) = f(b_-) = i$  (and forgetting the signs of the arrows).

Let  $\text{lbl}_n(D)$  be the set of  $n$ -labelings  $f$  of  $D$  such that  $f(\text{edg}(D)) = \{1, \dots, n\}$ ,  $|f| = n - 1$ , and the  $f$ -cutting arrows of  $D$  are pairwise unlinked (in the sense of Section 3.1). Then

$$(9.3.1) \quad \nabla([D]) = \sum_{n=1} \sum_{f \in \text{lbl}_n(D)} \frac{(-1)^{|f|} z^{n-1}}{n!} \prod_{i=1}^n \langle \underline{D}_{f,i} \rangle \in R[\mathcal{S}_0].$$

The expression on the right-hand side is finite since  $\text{lbl}_n(D) = \emptyset$  for  $n > \# \text{edg}(D)$ . The set  $\text{lbl}_1(D)$  consists of only one element  $f = 1$  so that the free term of  $\nabla([D])$  is  $\langle \underline{D} \rangle$ .

## 10. PROOF OF THEOREM 9.3.1

The proof of Theorem 9.3.1 largely follows the proof of Theorems 9.2 and 13.2 in [Tu2]. We therefore expose only the main lines of the proof. The key point behind Theorem 9.3.1 is the existence of a natural comultiplication in  $\mathcal{E}$  and we define it first. Then we construct  $\nabla$  and prove that it is an isomorphism.

**10.1. Comultiplication in  $\mathcal{E}$ .** We need to study more extensively the labelings of arrow diagrams defined at the end of the previous section. Let  $D$  be an arrow diagram with core circle  $S$ . Each  $n$ -labeling  $f$  of  $D$  gives rise to  $n$  monomials  $D_{f,1}, \dots, D_{f,n} \in \mathcal{E}$  as follows. Identifying  $a = b$  for every  $f$ -cutting arrow  $(a, b)$  of  $D$ , we transform  $S$  into a 4-valent graph,  $\Gamma^f$ , with  $|f|$  vertices. The projection  $S \rightarrow \Gamma^f$  maps the non- $f$ -cutting arrows of  $D$  into “arrows” on  $\Gamma^f$ , i.e., into ordered pairs of (distinct) generic points of  $\Gamma^f$ . The labeling  $f$  induces a labeling of the edges of  $\Gamma^f$  by the numbers  $1, 2, \dots, n$ . It follows from the definition of a labeling that for each  $i = 1, \dots, n$ , the union of edges of  $\Gamma^f$  labeled with  $i$  is a disjoint union of  $r_i = r_i(f) \geq 0$  circles  $S_1^i, \dots, S_{r_i}^i$ . The orientation of  $S$  induces an orientation of the edges of  $\Gamma^f$  and of these circles. We transform each circle  $S_j^i$  with  $j = 1, \dots, r_i$  into an arrow diagram by adding to it all the arrows of  $\Gamma^f$  with both endpoints on  $S_j^i$ . The signs of these arrows are by definition the signs of the corresponding non- $f$ -cutting arrows of  $D$ . Set

$$D_{f,i} = \prod_{j=1}^{r_i} [S_j^i] \in \mathcal{E}.$$

For any  $n \geq 2$ , denote  $\text{Lbl}_n(D)$  the set of  $n$ -labelings  $f$  of  $D$  such that the  $f$ -cutting arrows of  $D$  are pairwise unlinked. The latter condition can be reformulated in terms of the numbers  $r_1(f), \dots, r_n(f)$  introduced above:  $f \in \text{Lbl}_n(D)$  if and only if  $r_1(f) + \dots + r_n(f) = |f| + 1$ . For  $f \in \text{Lbl}_n(D)$ , set

$$\Delta(D, f) = (-1)^{|f|} z^{|f|} D_{f,1} \otimes D_{f,2} \otimes \dots \otimes D_{f,n} \in \mathcal{E}^{\otimes n}$$

where  $\mathcal{E}^{\otimes n}$  is the tensor product over  $R$  of  $n$  copies of  $\mathcal{E}$ .

By a comultiplication in  $\mathcal{E}$ , we mean a coassociative algebra homomorphism  $\Delta : \mathcal{E} \rightarrow \mathcal{E} \otimes \mathcal{E}$ . (The coassociativity means that  $(\text{id} \otimes \Delta)\Delta = (\Delta \otimes \text{id})\Delta$ .) We claim that the formula

$$\Delta([D]) = \sum_{f \in \text{Lbl}_2(D)} \Delta(D, f) \in \mathcal{E} \otimes \mathcal{E}$$

extends by multiplicativity to a well-defined comultiplication in  $\mathcal{E}$ . This can be deduced from [Tu2], Theorem 9.2 or proven directly repeating the same arguments. We explain how to deduce our claim from [Tu2]. Comparing the definition of  $\Delta([D])$  with the comultiplication in the algebra of skein classes of knots in  $(\text{a surface}) \times \mathbb{R}$  given in [Tu2], we observe that they correspond to each other provided  $D$  underlies a knot diagram on the surface. (The variables  $h = h_1, \bar{h} = h_{-1}$  used in [Tu2] should be replaced with 0 and  $z$ , respectively. After the substitution  $h = 0$ , we can consider only labelings satisfying - in the notation of [Tu2] - the condition  $\|f\| = -|f|$  which translates here as the assumption that the  $f$ -cutting arrows of  $D$  are pairwise unlinked.) The results of [Tu2] imply that if a move  $(a)_{ad}, (b)_{ad}, (c)_{ad}$  on  $D$  underlies a Reidemeister move on a knot diagram, then  $\Delta([D])$  is preserved under this move. Since any arrow diagram  $D$  underlies a knot diagram on a surface and any move  $(a)_{ad}, (b)_{ad}, (c)_{ad}$  on  $D$  can be induced by a Reidemeister move, we conclude that  $\Delta([D])$  is invariant under the moves  $(a)_{ad}, (b)_{ad}, (c)_{ad}$  on  $D$ . Therefore the formula  $[D] \mapsto \Delta([D])$  yields a well-defined mapping  $\mathcal{K} \rightarrow \mathcal{E} \otimes \mathcal{E}$ . This mapping uniquely extends to an algebra homomorphism  $R[\mathcal{K}] \rightarrow \mathcal{E} \otimes \mathcal{E}$ . The results of [Tu2] imply that for an arrow diagram  $D$  underlying a knot diagram on a surface and any arrow  $e$  of  $D$  with  $\text{sign}(e) = +$ , the skein relation  $[D] - [D_e^-] - z[D_e'] [D_e'']$  lies in the kernel of the latter homomorphism. The condition that  $D$  underlies a knot diagram is verified for all  $D$ . Therefore the



homomorphism  $R[\mathcal{K}] \rightarrow \mathcal{E} \otimes \mathcal{E}$  annihilates the skein ideal and induces an algebra homomorphism  $\Delta : \mathcal{E} \rightarrow \mathcal{E} \otimes \mathcal{E}$ . The coassociativity of  $\Delta$  follows from the easy formulas

$$(\text{id} \otimes \Delta)\Delta([D]) = \sum_{f \in \text{Lbl}_3(D)} \Delta(D, f) = (\Delta \otimes \text{id})\Delta([D])$$

(cf. [Tu2], p. 665). More generally, for any  $n \geq 2$ , the value on  $[D] \in \mathcal{E}$  of the iterated homomorphism

$$\Delta^{(n)} = (\text{id}^{\otimes(n-1)} \otimes \Delta) \circ (\text{id}^{\otimes(n-2)} \otimes \Delta) \circ \dots \circ (\text{id} \otimes \Delta) \Delta : \mathcal{E} \rightarrow \mathcal{E}^{\otimes(n+1)}$$

is computed by

$$\Delta^{(n)}([D]) = \sum_{f \in \text{Lbl}_{n+1}(D)} \Delta(D, f).$$

Note for the record that each arrow diagram  $D$  admits constant 2-labelings  $f_1, f_2$  taking values 1, 2 on all edges, respectively. The corresponding summands of  $\Delta([D])$  are  $\Delta(D, f_1) = [D] \otimes 1$  and  $\Delta(D, f_2) = 1 \otimes [D]$ .

**10.2. Homomorphism  $\nabla : \mathcal{E} \rightarrow R[\mathcal{S}_0]$ .** There are two obvious  $R$ -linear homomorphisms  $\varepsilon : \mathcal{E} \rightarrow R$  and  $q : \mathcal{E} \rightarrow R[\mathcal{S}_0]$ . The homomorphism  $\varepsilon$  sends  $1 \in \mathcal{E}$  into  $1 \in R$  and sends all virtual knots and their non-void products into 0. The homomorphism  $q$  sends 1 and all products of  $\geq 2$  virtual knots into 0 and sends a virtual knot  $K$  into  $\langle \underline{K} \rangle$ . Tensorizing  $q$  with itself, we obtain for all  $n \geq 1$  a homomorphism  $q^{\otimes n} : \mathcal{E}^{\otimes n} \rightarrow R[\mathcal{S}_0]^{\otimes n}$ . Let  $s_n : R[\mathcal{S}_0]^{\otimes n} \rightarrow R[\mathcal{S}_0]$  be the  $R$ -linear homomorphism sending  $a_1 \otimes \dots \otimes a_n$  into  $(n!)^{-1} a_1 \dots a_n$ . Set

$$\nabla = \varepsilon + q + \sum_{n \geq 2} s_n q^{\otimes n} \Delta^{(n-1)} : \mathcal{E} \rightarrow R[\mathcal{S}_0]$$

where  $\Delta^{(1)} = \Delta$ . It is clear that  $\nabla$  is  $R$ -linear. The same argument as in [Tu2], Lemma 13.4 shows that  $\nabla$  is an algebra homomorphism. Computing  $\nabla$  on the skein class of an arrow diagram  $D$ , we obtain

$$\nabla([D]) = \sum_{n \geq 1} s_n q^{\otimes n} \sum_{f \in \text{Lbl}_n(D)} \Delta(D, f) = \sum_{n \geq 1} \sum_{f \in \text{Lbl}_n(D)} \frac{(-1)^{|f| - z|f|}}{n!} \prod_{i=1}^n q(D_{f,i}).$$

Note that  $q(D_{f,i}) = 0$  unless  $r_i(f) = 1$  in which case  $q(D_{f,i}) = \langle \underline{D}_{f,i} \rangle$ . For a labeling  $f \in \text{Lbl}_n(D)$  the equalities  $r_1(f) = \dots = r_n(f) = 1$  are equivalent to the inclusion  $f \in \text{lbl}_n(D)$ . This yields Formula 9.3.1.

Observe that  $\nabla([D])$  is a sum of  $\langle \underline{D} \rangle$  and a polynomial in strings of rank  $< \text{rank } D$ . An induction on the rank of strings shows that the image of  $\nabla$  contains all strings. Therefore  $\nabla$  is surjective.

The proof of the injectivity of  $\nabla$  is based on the following lemma.

**Lemma 10.2.1.** *There is a  $\mathbb{Q}$ -valued function  $\eta$  on the set of isomorphism classes of (finite) oriented trees such that the following holds:*

- (i) *if  $T$  is a tree with one vertex and no edges, then  $\eta(T) = 1$ ;*
- (ii) *if an oriented tree  $T'$  (resp.  $U$ ) is obtained from an oriented tree  $T$  by reversing the orientation of an edge  $e$  (resp. by contracting  $e$  into a point), then  $\eta(T) + \eta(T') + \eta(U) = 0$ ;*
- (iii) *if an oriented tree  $T'$  (resp.  $T''$ ) is obtained from an oriented tree  $T$  by replacing two distinct edges with common origin  $ab, ac$  by  $ab, bc$  (resp. by  $ac, cb$ ) and if  $U$  is obtained from  $T$  by identifying  $b$  with  $c$  and  $ab$  with  $ac$ , then  $\eta(T) = \eta(T') + \eta(T'') + \eta(U)$ .*

In this lemma by an edge  $ab$  we mean an *oriented* edge directed from  $a$  to  $b$ .

Lemma 10.2.1 was first established in [Tu2], Theorem 14.1 where it is also shown that  $\eta$  is unique (we shall not need this). The construction in [Tu2] is indirect and does not provide an explicit formula for  $\eta$ . Such a formula was pointed out by François Jaeger [Ja]. The following proof of Lemma 10.2.1 is a simplified version of the proof given by Jaeger [Ja].

*Proof.* By a *forest* we shall mean a disjoint union of a finite family of finite oriented trees. The set of vertices of a forest  $F$  is denoted  $V(F)$ . For a forest  $F$  and an integer  $n \geq 1$ , denote by  $C_n(F)$  the set of surjective mappings  $f : V(F) \rightarrow \{1, \dots, n\}$  such that for every edge  $ab$  of  $F$  we have  $f(a) < f(b)$ . This set is empty for  $n > \#(V(F))$ . Set

$$\eta(F) = \sum_{n \geq 1} \frac{(-1)^{n+1}}{n} \#(C_n(F)) \in \mathbb{Q}.$$

We claim that  $\eta$  satisfies all the conditions of the lemma. Condition (i) is obvious. Condition (iii) is a direct corollary of the definitions. Indeed for all  $n$ , the set  $C_n(T)$  splits as a disjoint union of the sets

$C_n(T'), C_n(T''), C_n(U)$ . Hence  $\#(C_n(T)) = \#(C_n(T')) + \#(C_n(T'')) + \#(C_n(U))$  and  $\eta(T) = \eta(T') + \eta(T'') + \eta(U)$ . It remains to verify (ii). Let  $F$  be obtained from  $T$  by removing the interior of the edge  $e$ . For all  $n$ , the set  $C_n(F)$  splits as a disjoint union of the sets  $C_n(T), C_n(T'), C_n(U)$ . Hence  $\#(C_n(F)) = \#(C_n(T)) + \#(C_n(T')) + \#(C_n(U))$  and  $\eta(F) = \eta(T) + \eta(T') + \eta(U)$ . Thus we need only to prove that  $\eta(F) = 0$  for every forest  $F$  with two components  $T_1, T_2$ .

For non-negative integers  $n, k_1, k_2$ , denote by  $C_n(k_1, k_2)$  the set of pairs  $(l_1, l_2)$  where for  $i = 1, 2$ ,  $l_i$  is an order-preserving injection from  $\{1, \dots, k_i\}$  into  $\{1, \dots, n\}$  and  $l_1(\{1, \dots, k_1\}) \cup l_2(\{1, \dots, k_2\}) = \{1, \dots, n\}$ . Having  $g_1 \in C_{k_1}(T_1)$ ,  $g_2 \in C_{k_2}(T_2)$  and having  $(l_1, l_2) \in C_n(k_1, k_2)$  we define a mapping  $f = f(g_1, g_2, l_1, l_2) : V(F) \rightarrow \{1, \dots, n\}$  by  $f(v) = l_1 g_1(v)$  for  $v \in V(T_1)$  and  $f(v) = l_2 g_2(v)$  for  $v \in V(T_2)$ . Clearly,  $f \in C_n(F)$ . It is obvious that any  $f \in C_n(F)$  can be uniquely presented in the form  $f = f(g_1, g_2, l_1, l_2)$  where  $g_i \in C_{k_i}(T_i)$  with  $k_i = \#(f(V(T_i))) \geq 1$  for  $i = 1, 2$ . Therefore

$$\begin{aligned} \eta(F) &= \sum_{n \geq 1} \frac{(-1)^{n+1}}{n} \left( \sum_{k_1, k_2 \geq 1} \sum_{g_1 \in C_{k_1}(T_1), g_2 \in C_{k_2}(T_2)} \#(C_n(k_1, k_2)) \right) \\ &= \sum_{k_1, k_2 \geq 1} \#(C_{k_1}(T_1)) \#(C_{k_2}(T_2)) \sum_{n \geq 1} \frac{(-1)^{n+1}}{n} \#(C_n(k_1, k_2)). \end{aligned}$$

Thus it is enough to prove that for all  $k_1 \geq 1, k_2 \geq 1$ , the numbers  $c_n(k_1, k_2) = \#(C_n(k_1, k_2))$  verify

$$(10.2.1) \quad \sum_{n \geq 1} \frac{(-1)^{n+1}}{n} c_n(k_1, k_2) = 0.$$

Clearly,  $c_n(k_1, k_2)$  is the number of pairs  $(S_1, S_2)$  where  $S_1, S_2$  are subsets of  $\{1, \dots, n\}$  such that  $S_1 \cup S_2 = \{1, \dots, n\}$ ,  $\#(S_1) = k_1$ ,  $\#(S_2) = k_2$ . In particular,  $c_n(k_1, k_2) = 0$  if  $k_1 + k_2 < n$  or  $k_1 > n$  or  $k_2 > n$ . For any  $n \geq 1$  and commuting variables  $x, y$ ,

$$(x + y + xy)^n = \sum_{k_1, k_2 \geq 0} c_n(k_1, k_2) x^{k_1} y^{k_2}.$$

Therefore

$$\begin{aligned} \log(1 + x + y + xy) &= \sum_{n \geq 1} \frac{(-1)^{n+1}}{n} (x + y + xy)^n \\ &= \sum_{n \geq 1} \sum_{k_1, k_2 \geq 0} \frac{(-1)^{n+1}}{n} c_n(k_1, k_2) x^{k_1} y^{k_2} = \sum_{k_1, k_2 \geq 0} \left( \sum_{n \geq 1} \frac{(-1)^{n+1}}{n} c_n(k_1, k_2) \right) x^{k_1} y^{k_2}. \end{aligned}$$

Since

$$\log(1 + x + y + xy) = \log((1 + x)(1 + y)) = \log(1 + x) + \log(1 + y),$$

the terms with  $k_1 \geq 1, k_2 \geq 1$  in the above series must vanish. This gives Formula 10.2.1.  $\square$

**10.3. The injectivity of  $\nabla : \mathcal{E} \rightarrow R[S_0]$ .** We begin by associating with any virtual string  $\alpha$  an element  $\zeta(\alpha) \in \mathcal{E}$ . Let  $S$  be the core circle of  $\alpha$ . A *surgery* along an arrow  $(a, b) \in \text{arr}(\alpha)$  consists in picking two (positively oriented) arcs  $aa^+, bb^+ \subset S$  and then quotienting the complement of their interiors  $S - ((aa^+)^\circ \cup (bb^+)^\circ)$  by  $a = b^+, b = a^+$ . It is understood that the arcs  $aa^+, bb^+$  are small enough not to contain endpoints of  $\alpha$  besides  $a, b$ , respectively. Such a surgery transforms  $S$  into two disjoint oriented circles. We make each of them into a string by adding all the arrows of  $\alpha$  with both endpoints on the arc  $a^+b$  (resp. on  $ba^+$ ). (The arrows of  $\alpha$  with one endpoint on  $ab$  and the other one on  $ba$  disappear under surgery.)

Let us call a set  $F \subset \text{arr}(\alpha)$  *special* if the arrows of  $\alpha$  belonging to  $F$  are pairwise unlinked. Applying surgery inductively to all arrows of  $\alpha$  belonging to a special set  $F$ , we transform  $\alpha$  into  $n = \#(F) + 1$  strings. Providing all the arrows of these strings with sign  $+$ , we obtain  $n$  arrow diagrams  $D_1^F, \dots, D_n^F$ . Note that they have together at most  $\#(\text{arr}(\alpha)) - \#(F)$  arrows. We now define an oriented graph  $\Gamma_F$ . The vertices of  $\Gamma_F$  are the symbols  $v_1, \dots, v_n$ . Two vertices  $v_i, v_j$  are related by an oriented edge leading from  $v_i$  to  $v_j$  if there is an arrow  $(a, b) \in F$  such that the arcs  $aa^+, bb^+ \subset S$  involved in the surgery along this arrow lie on the core circles of  $D_i^F, D_j^F$ , respectively. It is easy to see that  $\Gamma_F$  is a tree. Set

$$\zeta(\alpha) = \sum_{F \subset \text{arr}(\alpha)} \eta(\Gamma_F) z^{\#(F)} \prod_{i=1}^{\#(F)+1} [D_i^F] \in \mathcal{E}$$

where  $F$  runs over all special subsets of  $\text{arr}(\alpha)$ . The summand corresponding to  $F = \emptyset$  is the string  $\alpha$  itself with sign  $+$  on all arrows.

The key property of  $\zeta(\alpha) \in \mathcal{E}$  is its invariance under the basic homotopy moves on  $\alpha$ . This follows from [Tu2], Lemma 15.1.1 in the case where the moves are realized geometrically by homotopy of a curve realizing  $\alpha$  on a surface. Since the homotopy moves can be always realized geometrically, the result follows. The mapping  $\alpha \mapsto \zeta(\alpha)$  extends by multiplicativity to an algebra homomorphism  $R[S_0] \rightarrow \mathcal{E}$  denoted also  $\zeta$ .

We can now prove the injectivity of  $\nabla$ . For  $r \geq 0$ , denote by  $B_r$  the  $R$ -submodule of  $\mathcal{E}$  additively generated by monomials  $[D_1][D_2] \cdots [D_n]$  such that the total number of arrows in the arrow diagrams  $D_1, D_2, \dots, D_n$  is less than or equal to  $r$ . Clearly,  $0 = B_0 \subset B_1 \subset \dots$  and  $\cup_r B_r = \mathcal{E}$ . Pick  $b = [D_1][D_2] \cdots [D_n] \in B_r$ . Using the skein relation in  $\mathcal{E}$  it is easy to see that  $b \pmod{B_{r-1}} \in B_r/B_{r-1}$  does not depend on the signs of the arrows of  $D_1, \dots, D_n$ . This observation, Formula 9.3.1 and the definition of  $\zeta$  imply that  $(\zeta \nabla)(b) - b \in B_{r-1}$ . Therefore  $(\zeta \nabla - \text{id})^r(b) = 0$ . The inclusion  $b \in \text{Ker } \nabla$  would imply  $b = 0$ . Thus  $B_r \cap \text{Ker } \nabla = 0$ . Since  $\cup_r B_r = \mathcal{E}$ , we obtain  $\text{Ker } \nabla = 0$ .

**10.4. Remarks.** The comultiplication  $\Delta$  defined in Section 10.1 makes  $\mathcal{E}$  into a Hopf algebra over  $R$ . Its counit is the homomorphism  $\varepsilon : \mathcal{E} \rightarrow R$  used in the definition of  $\nabla$ . For an arrow diagram  $D$ , denote by  $\tilde{D}$  the same diagram with opposite signs on all arrows. The transformation  $[D] \mapsto -[\tilde{D}]$  preserves the skein relation and therefore induces an algebra automorphism of  $\mathcal{E}$ . This automorphism is an antipode for  $\mathcal{E}$ . This follows from the corresponding theorem for the skein algebras of curves on surfaces conjectured in [Tu2] and proven in [CR] and independently in [Pr]. In the construction of the Hopf algebra  $\mathcal{E}$  instead of the ground ring  $R = \mathbb{Q}[z]$  we can use  $\mathbb{Z}[z]$ . It is only to construct the homomorphisms  $\nabla$  and  $\zeta$  that we need  $\mathbb{Q}$ .

## 11. ALGEBRAS AND GROUPS ASSOCIATED WITH STRINGS

We discuss here various algebraic structures associated with closed and open strings. We begin by recalling the notion of a spiral Lie coalgebra and several related definitions from [Tu2], Section 11. Throughout the section,  $R$  is a commutative ring with unit and  $\otimes = \otimes_R$ .

**11.1. Spiral Lie coalgebras.** For a Lie coalgebra  $(A, \nu : A \rightarrow A^{\otimes 2})$  over  $R$  and an integer  $n \geq 1$ , set

$$\nu^{(n)} = (\text{id}_A^{\otimes(n-1)} \otimes \nu) \circ \cdots \circ (\text{id}_A \otimes \nu) \circ \nu : A \rightarrow A^{\otimes(n+1)}.$$

In particular,  $\nu^{(1)} = \nu$ . A Lie coalgebra  $(A, \nu)$  over  $R$  is *spiral*, if  $A$  is free as the  $R$ -module and the filtration  $\text{Ker } \nu^{(1)} \subset \text{Ker } \nu^{(2)} \subset \cdots$  exhausts  $A$ , i.e.,  $A = \cup_{n \geq 1} \text{Ker } \nu^{(n)}$ .

Assume from now on that  $A$  is spiral. The dual Lie algebra  $A^* = \text{Hom}_R(A, R)$  has the following completeness property. Consider the lower central series  $A^* = A^{*(1)} \supset A^{*(2)} \supset \cdots$  of  $A^*$  where  $A^{*(n+1)} = [A^{*(n)}, A^*]$  for  $n \geq 1$ . Let  $a_1, a_2, \dots \in A^*$  be an infinite sequence such that for any  $n \geq 1$  all terms of the sequence starting from a certain place belong to  $A^{*(n)}$ . Clearly, if  $x \in \text{Ker } \nu^{(n)}$  and  $a \in A^{*(n+1)}$ , then  $a(x) = 0$ . Since  $A = \cup_n \text{Ker } \nu^{(n)}$ , the sum  $a(x) = a_1(x) + a_2(x) + \cdots$  contains only a finite number of non-zero terms for every  $x \in A$ . Therefore  $a(x)$  is a well-defined element of  $R$ . The formula  $x \mapsto a(x) : A \rightarrow R$  defines an element of  $A^*$  denoted  $a_1 + a_2 + \cdots$  and called the (infinite) sum of  $a_1, a_2, \dots$ . A similar argument shows that  $\cap_n A^{*(n)} = 0$  and the natural Lie algebra homomorphism  $A^* \rightarrow \text{projlim}_n (A^*/A^{*(n)})$  is an isomorphism.

For  $a, b \in A^*$ , consider the sum

$$\mu(a, b) = a + b + \frac{1}{2}[a, b] + \frac{1}{12}([a, [a, b]] + [b, [b, a]]) + \cdots \in A^*$$

where the right-hand side is the Campbell-Hausdorff series for  $\log(e^a e^b)$ , see [Se]. The resulting mapping  $\mu : A^* \times A^* \rightarrow A^*$  is a group multiplication in  $A^*$ . Here  $a^{-1} = -a$  and  $0$  is the group unit. The group  $(A^*, \mu)$  is denoted  $\text{Exp } A^*$ . Heuristically, this is the ‘‘Lie group’’ with Lie algebra  $A^*$ . The equality  $A^* = \text{projlim}_n (A^*/A^{*(n)})$  implies that the group  $\text{Exp } A^*$  is pro-nilpotent.

Consider the symmetric (commutative and associative) algebra of  $A$ :

$$S = S(A) = \oplus_{n \geq 0} S^n(A).$$

Here  $S^0(A) = R$ ,  $S^1(A) = A$ , and  $S^n(A)$  is the  $n$ -th symmetric tensor power of  $A$  for  $n \geq 2$ . The unit  $1 \in R = S^0(A)$  is the unit of  $S$ . The group multiplication  $\mu : A^* \times A^* \rightarrow A^*$  induces a comultiplication  $S \rightarrow S \otimes S$  as follows. Since  $A$  is a free  $R$ -module, the natural map  $A \rightarrow (A^*)^*$  extends to an embedding of  $S$  into the algebra of  $R$ -valued functions on  $A^*$ . We can identify  $S$  with the image of this embedding. Similarly, we can identify  $S \otimes S$  with an algebra of  $R$ -valued functions on  $A^* \times A^*$ . It is easy to observe that

for any  $x \in S$ , we have  $x \circ \mu \in S \otimes S$ . Indeed, it suffices to prove this for  $x \in A$ . Then  $x \in \text{Ker } \nu^{(n)}$  for some  $n$  so that  $x$  annihilates all but finite number of terms of the Campbell-Hausdorff series. Our claim follows then from the duality between the Lie bracket in  $A^*$  and the Lie cobracket  $\nu$ . For example, if  $n = 3$  and  $\nu^{(2)}(x) = \sum_i \alpha_i \otimes \beta_i \otimes \gamma_i \in A^{\otimes 3}$ , then

$$x \circ \mu = x \otimes 1 + 1 \otimes x + \frac{1}{2}\nu(x) + \frac{1}{12} \sum_i (\alpha_i \beta_i \otimes \gamma_i + \gamma_i \otimes \alpha_i \beta_i).$$

The formula  $\Delta(x) = x \circ \mu$  defines a coassociative comultiplication in  $S$ . It has a counit  $S \rightarrow R$  defined as the projection to  $S^0(A) = R$ . The antipode  $S \rightarrow S$  is the algebra homomorphism sending any  $x \in A$  into  $-x \in A$ . A routine check shows that  $S$  is a (commutative) Hopf algebra. Heuristically, it should be viewed as the Hopf algebra of  $R$ -valued functions on the group  $\text{Exp } A^*$  or as the Hopf dual of the universal enveloping algebra of  $A^*$ .

The construction of  $\text{Exp } A^*$  and  $S(A)$  can be generalized as follows. Pick  $h \in R$  and observe that the mapping  $h\nu : A \rightarrow A \otimes A$  is a Lie cobracket in  $A$ . It induces the Lie bracket  $[\cdot, \cdot]_h = h[\cdot, \cdot]$  in  $A^*$  where  $[\cdot, \cdot]$  is the Lie bracket induced by  $\nu$ . The corresponding multiplication  $\mu_h$  in  $A^*$  is given by

$$\mu_h(a, b) = a + b + \frac{h}{2}[a, b] + \frac{h^2}{12}([a, [a, b]] + [b, [b, a]]) + \dots$$

This multiplication makes  $A^*$  into a group denoted  $\text{Exp}_h A^*$ . As above,  $\mu_h$  induces a comultiplication in the symmetric algebra  $S = S(A)$ . This makes  $S$  into a Hopf algebra over  $R$  denoted  $S_h(A)$ . For  $h = 1$ , we obtain the same objects as in the previous paragraphs. Note for the record that for any  $h \in R$ , the formula  $a \mapsto ha : A^* \rightarrow A^*$  defines a group homomorphism  $\text{Exp}_h A^* \rightarrow \text{Exp } A^*$ . If  $h \in R$  is a non-zero-divisor, this homomorphism is injective.

It is clear that the construction of  $\text{Exp}_h A^*$  and  $S_h(A)$  is functorial. For a Lie coalgebra homomorphism  $\psi$  from  $A$  into a spiral Lie coalgebra  $A'$ , the dual homomorphism  $\psi^* : (A')^* \rightarrow A^*$  preserves the Lie bracket  $[\cdot, \cdot]_h$  and the group multiplication  $\mu_h$ . The algebra homomorphism  $S_h(A) \rightarrow S_h(A')$  induced by  $\psi$  is a homomorphism of Hopf algebras.

**Lemma 11.1.1.** *The Lie coalgebra of strings  $(\mathcal{A}_0 = \mathcal{A}_0(R), \nu)$  defined in Section 8.2 is spiral.*

*Proof.* This follows from the obvious fact that  $\nu^{(n)}(\langle \alpha \rangle) = 0$  for any string  $\alpha$  of rank  $\leq n$ . (Actually a stronger assertion holds:  $\nu^{(n)}(\langle \alpha \rangle) = 0$  for any string  $\alpha$  of rank  $\leq 4n + 2$ .)  $\square$

Applying the constructions above to the Lie coalgebra  $\mathcal{A}_0$  and any  $h \in R$ , we obtain a group  $\text{Exp}_h \mathcal{A}_0^*$  and a Hopf algebra  $S_h(\mathcal{A}_0)$  over  $R$ . Note that  $S_h(\mathcal{A}_0) = R[S_0]$  as algebras.

**Theorem 11.1.2.** *For  $R = \mathbb{Q}[z]$  and  $h = z \in R$ , the homomorphism  $\nabla : \mathcal{E} \rightarrow R[S_0] = S_h(\mathcal{A}_0)$  is an isomorphism of Hopf algebras.*

The proof of this theorem follows the lines of [Tu2], Section 12 and Lemma 13.5; we omit the details.

**11.2. Remarks.** 1. The equality  $\mathcal{A} = \mathcal{A}_0 \oplus R$  implies that  $\text{Exp } \mathcal{A}^* = \text{Exp } \mathcal{A}_0^* \times \underline{R}$  where  $\mathcal{A}_0^* = (\mathcal{A}_0)^*$  and  $\underline{R}$  is the additive group of  $R$ . More generally, for any  $h \in R$ , we have  $\text{Exp}_h \mathcal{A}^* = \text{Exp}_h \mathcal{A}_0^* \times \underline{R}$ .

2. For any  $h \in R$ , the Lie coalgebra  $\mathcal{A}_{r,g}$  defined in Section 8.5 gives rise to a group  $\text{Exp}_h \mathcal{A}_{r,g}^*$  and a Hopf algebra  $S_h(\mathcal{A}_{r,g})$  which are quotients of  $\text{Exp}_h \mathcal{A}_0^*$  and  $S_h(\mathcal{A}_0)$ , respectively. The Lie coalgebra  $Z = Z(R)$  discussed in Section 8.6 is known to be spiral, so that we have the associated group  $\text{Exp}_h Z^*$  and the Hopf algebra  $S_h(Z)$ . The homomorphism  $\psi_0 : Z \rightarrow \mathcal{A}_0$  induces a group homomorphism  $\text{Exp}_h \mathcal{A}_0^* \rightarrow \text{Exp}_h Z^*$  and a Hopf algebra homomorphism  $S_h(Z) \rightarrow S_h(\mathcal{A}_0)$ .

**11.3. The algebra of open strings.** We begin with algebraic preliminaries. Recall that a *module* over a Lie algebra  $(L, [\cdot, \cdot] : L^{\otimes 2} \rightarrow L)$  over  $R$  can be defined as an  $R$ -module  $M$  endowed with an  $R$ -linear homomorphism  $\rho : L \otimes M \rightarrow M$  such that

$$(11.3.1) \quad \rho([\cdot, \cdot] \otimes \text{id}_M) = \rho(\text{id}_L \otimes \rho)(\text{id}_{L \otimes L \otimes M} - \text{Perm}_L \otimes \text{id}_M) : L \otimes L \otimes M \rightarrow M$$

where  $\text{Perm}_L$  is the permutation  $x \otimes y \mapsto y \otimes x$  in  $L^{\otimes 2} = L \otimes L$ . (Formula 11.3.1 is equivalent to the usual identity  $[x, y]m = x(ym) - y(xm)$  for  $x, y \in L, m \in M$ .) Dually, a *comodule* over a Lie coalgebra  $(A, \nu : A \rightarrow A^{\otimes 2})$  over  $R$  is an  $R$ -module  $M$  endowed with an  $R$ -linear homomorphism  $\rho : M \rightarrow A \otimes M$  such that

$$(11.3.2) \quad (\nu \otimes \text{id}_M)\rho = (\text{id}_{A \otimes A \otimes M} - \text{Perm}_A \otimes \text{id}_M)(\text{id}_A \otimes \rho)\rho : M \rightarrow A \otimes A \otimes M.$$

Such  $M$  is automatically a module over the dual Lie algebra  $A^* = \text{Hom}_R(M, R)$ : an element  $a \in A^*$  acts on  $M$  by the endomorphism  $\varphi_a : M \rightarrow M$  sending  $m \in M$  to  $-(a \otimes \text{id}_M)\rho(m) \in R \otimes M = M$ .

A comodule  $(M, \rho)$  over a Lie coalgebra  $A$  is *spiral* if  $M = \bigcup_{n \geq 1} \text{Ker } \rho^{(n)}$  where

$$\rho^{(n)} = (\text{id}_A^{\otimes(n-1)} \otimes \rho) \circ \cdots \circ (\text{id}_A \otimes \rho) \circ \rho : M \rightarrow A^{\otimes n} \otimes M.$$

If both  $A$  and  $M$  are spiral, then the action of  $A^*$  on  $M$  integrates into a group action of the group  $\text{Exp } A^*$  on  $M$  defined by

$$am = e^{\varphi_a}(m) = m + \sum_{k \geq 1} (\varphi_a)^k(m)/k!$$

for  $a \in \text{Exp } A^* = A^*$ ,  $m \in M$ . Note that for  $m \in \text{Ker } \rho^{(n)}$  the sum on the right-hand side has at most  $n$  non-zero terms.

Let  $\mathcal{M} = \mathcal{M}(R)$  be the free  $R$ -module freely generated by the set of homotopy classes of open virtual strings. We provide  $\mathcal{M}$  with the structure of a comodule over the Lie coalgebra  $\mathcal{A}_0$ . Let  $\langle \beta \rangle$  be the generator of  $\mathcal{M}$  represented by an open string  $\beta$ . For an arrow  $e \in \text{arr}(\beta)$ , a surgery along  $e$  defined as in Section 10.3 transforms  $\beta$  into a disjoint union of a closed string  $\alpha_e$  and an open string  $\beta_e$ . Set

$$\rho(\langle \beta \rangle) = \sum_{e \in \text{arr}_+(\beta)} \langle \alpha_e \rangle \otimes \langle \beta_e \rangle - \sum_{e \in \text{arr}_-(\beta)} \langle \alpha_e \rangle \otimes \langle \beta_e \rangle \in \mathcal{A}_0 \otimes \mathcal{M}.$$

A direct computation shows that this gives a well-defined  $R$ -linear homomorphism  $\rho : \mathcal{M} \rightarrow \mathcal{A}_0 \otimes \mathcal{M}$  satisfying Formula 11.3.2. Thus  $\mathcal{M}$  is a comodule over  $\mathcal{A}_0$ . Combining  $\rho$  with the inclusion  $\mathcal{A}_0 \subset \mathcal{A}$  we obtain that  $\mathcal{M}$  is a comodule over  $\mathcal{A}$  as well. It is easy to see that  $\mathcal{M}$  is spiral. The construction above gives a group action of  $\text{Exp } \mathcal{A}^*$  on  $\mathcal{M}$ .

**11.4. Exercises.** 1. The obvious multiplication of open strings makes  $\mathcal{M}$  into an associative algebra with unit. Check that the group  $\text{Exp } \mathcal{A}^*$  acts on  $\mathcal{M}$  by algebra automorphisms.

2. Let  $cl : \mathcal{M} \rightarrow \mathcal{A}$  be the  $R$ -linear homomorphism induced by closing open strings. Check that for any open string  $\beta$ , we have  $\nu(\langle cl(\beta) \rangle) = (\text{id}_{\mathcal{A} \otimes \mathcal{A}} - \text{Perm}_{\mathcal{A}})(\text{id}_{\mathcal{A}} \otimes cl)\rho(\langle \beta \rangle)$ .

## 12. OPEN QUESTIONS

1. Find out whether the slice genus of a string is a homotopy invariant. The slice genus is invariant under  $(c)_s$  but may possibly decrease under  $(a)_s$ ,  $(b)_s$ . Find out whether these moves can transform a non-slice string into a slice one.

2. Find out which primitive based matrices  $T_0$  can be realized as  $T_0(\alpha)$  for a string  $\alpha$ . A necessary condition pointed out in Section 3.2 says that  $u'_{T_0}(1) = 0$ . Are there other conditions? Note that for the based matrix  $T(\alpha) = (G, s, b)$ , we have  $|b(e, f)| \leq \#(G) - 2$  for all  $e, f \in G$ . This however yields no conditions on the primitive based matrices arising from strings, since such a matrix  $T_0 = (G_0, s, b_0)$  may arise from a string of a rank much bigger than  $\#(G_0)$ .

3. Find further obstructions to the sliceness of a string. Specifically, are the virtual strings  $\alpha_{p,p}$  with  $p \geq 2$  slice (cf. Corollary 3.5.2)? The based matrix of  $\alpha_{p,p}$  is hyperbolic and gives no information on the question.

4. Consider the string of rank four  $\alpha = \alpha_\sigma$  where  $\sigma = (1342)$ . A direct computation shows that its primitive based matrix is trivial. Also  $\nu(\langle \alpha \rangle) = 0$  since  $\alpha$  has only 4 arrows. Is  $\alpha$  homotopically trivial? A more ambitious program would be to classify all strings of small rank (say,  $\leq 6$ ) up to homotopy.

5. Is multiplication of open strings discussed in Section 2.7 commutative (up to homotopy)? In other words, is the algebra of open strings  $\mathcal{M}$  considered at the end of Section 11 commutative?

6. Is there an invariant of formal knots combining the skein invariant  $\nabla$  with the Kontsevich universal finite type invariant of knots? This might lead to mixed arrow-chord diagrams.

7. Study invariants of virtual strings that change in a controlled way (say by constants) under the moves  $(a)_s$ ,  $(b)_s$ ,  $(c)_s$ , cf. the theory of Arnold's invariants of plane curves.

## REFERENCES

- [Ca] J. S. Carter, *Classifying immersed curves*, Proc. Amer. Math. Soc., 111 (1991), no. 1, 281–287.
- [CW] N. Chaves and C. Weber, *Plombages de rubans et problème des mots de Gauss*, Exposition. Math. 12 (1994), no. 1, 53–77 and no. 2, 124.
- [CR] P. Cotta-Ramusino, M. Rinaldi, *On the algebraic structure of link-diagrams on a 2-dimensional surface*, Comm. Math. Phys. 138 (1991), no. 1, 137–173.
- [DT] C. H. Dowker and M.B. Thistlethwaite, *Classification of knot projections*, Topology Appl. 16 (1983), no. 1, 19–31.

- [Fr] J. K. Francis, *The folded ribbon theorem. A contribution to the study of immersed circles*, Trans. Amer. Math. Soc. 141 (1969), 271–303.
- [Ga] C. F. Gauss, *Werke*, Vol. VIII, Teubner, Leipzig, 1900, pp. 272, 282–286.
- [GPV] M. Goussarov, M. Polyak, and O. Viro, *Finite-type invariants of classical and virtual knots*, Topology 39 (2000), no.5, 1045–1068.
- [Ja] F. Jaeger, *letter to the author*, 21.11.1991.
- [Ka] L. Kauffman, *Virtual knots theory*, European J. Combin. **20** (1999), no. 7, 663–690. Also in **math.GT/9811028**.
- [Pr] J. Przytycki, *Quantum group of links in a handlebody*, Deformation theory and quantum groups with applications to mathematical physics (Amherst, MA, 1990), 235–245, Contemp. Math., 134, Amer. Math. Soc., Providence, RI, 1992.
- [Ro] P. Rosenstiehl, *Solution algébrique du problème de Gauss sur la permutation des points d'intersection d'une ou plusieurs courbes fermées du plan*, C. R. Acad. Sci. Paris Sér. A-B 283 (1976), no. 8, A551–A553.
- [RR] P. Rosenstiehl and R. C. Read, *On the principal edge tripartition of a graph*, Advances in graph theory (Cambridge Combinatorial Conf., Trinity College, Cambridge, 1977). Ann. Discrete Math. 3 (1978), 195–226.
- [Se] J.-P. Serre, *Lie algebras and Lie groups*. Second edition. Lecture Notes in Mathematics, 1500. Springer-Verlag, Berlin, 1992.
- [Tu1] V. Turaev, *The Yang-Baxter equation and invariants of links*, Invent. Math. 92 (1988), no. 3, 527–553.
- [Tu2] V. Turaev, *Skein quantization of Poisson algebras of loops on surfaces*, Ann. Sci. École Norm. Sup. (4) 24 (1991), no. 6, 635–704.

IRMA, UNIVERSITÉ LOUIS PASTEUR - C.N.R.S.,  
 7 RUE RENÉ DESCARTES  
 F-67084 STRASBOURG  
 FRANCE